

PROCESSING AND ANTIGENICITY OF TAG-LINKED
GLYCOPROTEINS EXPRESSED IN MAMMALIAN
CELLS

Mark O'Reilly

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1997

Full metadata for this item is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/14512>

This item is protected by original copyright



Processing and Antigenicity of Tag-linked Glycoproteins Expressed in Mammalian Cells.

by

Mark O'Reilly BSc.

Div. of Cell & Molecular Biology

School of Medical and Biological Sciences

University of St. Andrews

A thesis submitted in partial fulfilment
of the requirements for
the degree of Doctor of Philosophy

August 1996



ProQuest Number: 10170853

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10170853

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TL
C 196

DECLARATION

I, Mark O'Reilly, hereby certify that this thesis, which is approximately 60,000 words in length, has been composed by myself, that it is a record of my own work and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed

.... Date 11.8.96

I was admitted as a research student in October 1992 and as a candidate for the degree of Doctor of Philosophy in October 1992; the higher study for which this is a record was carried out in the Faculty of Science of the University of St. Andrews between 1992 and 1996

Signed

. Date 11.8.96

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the Degree of Doctor of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Signed

.... Date 5/9/96

In submitting this thesis to the University of St. Andrews, I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Rick Randall for his help during the work, with special thanks going to Tomas Hanke, Dan Young and Bernie Precious for their technical advice and assistance, and to the remaining members of the group, recent and departed, for their helpful suggestions. I am indebted to the technicians in the building, especially to Bill Blythe for his photographic skills. Finally, I thank Jane, Alan, Jamsey, and my family for their support.

This work was supported by a grant from the AIDS Directed Programme of the Medical Research Council, U.K.

LIST OF ABBREVIATIONS

A	adenine
A#	absorbance at # nm
ABC	ATP-binding cassette
ADCC	antibody-dependent cell-mediated cytotoxicity
AIDS	acquired immunodeficiency syndrome
amp	ampicillin
APS	ammonium persulphate
ATP	adenosine 5'-tris(phosphate)
β-gal	β-galactosidase
β-mer	β-mercaptoethanol
BS ³	Bis(sulphosuccinimidyl) Suberate
BSA	bovine serum albumin
C	cytosine
CD#	cluster designation
cDNA	complementary DNA
cH	chimeric cheavy chain of antibody
C _H #	constant domain of an antibody heavy chain
CMI	cell-mediated immunity
CMV	Cytomegalovirus
CMV IE	CMV immediate-early enhancer/promoter
ConA	concanavalin A
cL	chimeric light chain of antibody
cpm	counts per minute
C-term.	carboxy terminus
CTL	cytotoxic T lymphocyte
DAPI	4, 6 diamidino-2-phenylindole
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
DEAE	diethylaminoethyl
dGTP	2'-deoxyguanosine 5'- triphosphate
DMF	dimethyl formamide
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNP	dinitrophenol
dNTP(s)	2'-deoxynucleotide 5'-triphosphate(s)
DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
EBNA-1	Epstein-Barr Virus nuclear antigen 1
EBV	Epstein-Barr Virus
ECL	Enhanced Chemiluminesence
<i>E. coli</i>	<i>Eschericia coli</i>
EDTA	ethylenediaminetetracetic acid
EGTA	ethyleneglycol-bis(b-aminoethylether) <i>N,N,N',N'</i> -tetraacetic acid

ELISA	enzyme-linked immunosorbent assay
EMCV	Encephalomyocarditis virus
F	fusion protein
Fab	antigen-binding fragment of antibody
Fc	crystalizable fragment of antibody
FCS	foetal calf serum FMDV Foot and mouth disease virus
G	guanine
G418	Geneticin 418
GMEM	Glasgow's modified essential medium
GMEM-10	Glasgow's modified essential medium plus 10% FCS
GOI	gene of interest.
HA	Haemagglutinin
hCMV	human cytomegalovirus
<i>his D</i>	Histidinol dehydrogenase
HisD	Histidinol dihydrochloride
HIV	human immunodeficiency virus
HN	Haemagglutinin-Neuraminidase
hPIV	human parainfluenza virus
HRP	Horseradish Peroxidase
HSV	Herpes Simplex Virus
<i>hyg B</i>	Hygromycin B phosphotransferase
HygB	Hygromycin B
Ig	immunoglobulin
IPTG	isopropyl- β -D -thiogalactopyranoside
KAc	potassium acetate
kb	kilobases
kbp	kilobase pairs
LB	Luria broth
LPS	lipopolysaccharide
LTAg	Large T Antigen
LTR	long terminal repeat
mAb	monoclonal antibody
MHC	major histocompatibility complex
MMTV	murine mammary tumor virus
moi	multiplicity of infection
mRNA	messenger RNA
MuV	Mumps Virus
MeV	Measles Virus
NaAc	sodium acetate
<i>neo apt</i>	Neomycin aminoglycoside phosphotransferase
NDV	Newcastle disease virus
NP	nucleoprotein
NP40	nonidet P-40

N-term.	Amino terminus
NTP(s)	ribonucleoside triphosphate(s)
OD#	optical density at # nm
ORF	open reading frame
<i>ori</i>	origin of replication
PAGE	polyacrylamide gel electrophoresis
PAGel	polyacrylamide gel
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque-forming units
pH	pondus hydrogen ($-\log_{10}[\text{H}^+]$)
PIV	Parainfluenza Virus
PMSF	phenylmethylsulphonylfluoride
RE	restriction endonuclease
rer	rough endoplasmic reticulum
RNA	ribonucleic acid
rpm	revolution per minute
RSV	Rous Sarcoma Virus
RSV LTR	Rous Sarcoma Virus Long Terminal Repeat
RT	room temperature
rTetR	reverse tetracycline repressor protein
rtTa	reverse tetracycline transactivator fusion protein
SAE	soluble antigen extract
SDS	sodium dodecyl sulphate
ser	smooth endoplasmic reticulum
SIV	simian immunodeficiency virus
SMAA	solid matrix-antibody-antigen complexes
SSC	salt-sodium citrate buffer
ssDNA	single-stranded DNA
strep	streptomycin
SV5	Simian Virus 5
SV40	Simian Virus 40
T	thymine
T D	T RNA Dpendent RNA Polymerase
T7-RDRP	T7 RNA Dpendent RNA Polymerase
TAE	Tris-acetate-EDTA buffer
TAP	transporter associated with antigen presentation
TAR	sequence at the 5' end of RNA recognized by tat and cellular proteins
TBE	Tris-borate-EDTA buffer
TCR	T cell receptor
TE	Tris-EDTA buffer
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
Tet ^c	tetracycline hydrochloride
TetR	tetracycline repressor protein
Tris	2-amino-2-(hydroxymethyl)propane-1,3-diol
TRP	tetracycline responsive promoter
tRNA	transfer RNA
tTa	tetracycline transactivator fusion protein

UV	ultra-violet
V _H	variable region of an antibody heavy chain
V _L	variable region of an antibody light chain
v/v	volume per volume ratio
WCE	whole cell extract
w/v	weight per volume ratio
X-gal	5-bromo-4-chloro-3-indolyl- β -D -galactoside

ABBREVIATIONS FOR AMINO ACIDS

Alanine	ala	A
Arginine	arg	R
Asparagine	asn	N
Aspartic acid	asp	D
Cysteine	cys	C
Glutamine	gln	Q
Glutamic acid	glu	E
Glycine	gly	G
Histidine	his	H
Isoleucine	ile	I
Leucine	leu	L
Lysine	lys	K
Methionine	met	M
Phenylalanine	phe	F
Proline	pro	P
Serine	ser	S
Threonine	thr	T
Tryptophan	try	W
Tyrosine	tyr	Y
Valine	val	V

GENETIC CODE

TTT phe F	TCT ser S	TAT tyr Y	TGT cys C
TTC phe F	TCC ser S	TAC tyr Y	TGC cys C
TTA leu L	TCA ser S	TAA OCH Z	TGA OPA Z
TTG leu L	TCG ser S	TAG AMB Z	TGG trp W

CTT leu L	CCT pro P	CAT his H	CGT arg R
CTC leu L	CCC pro P	CAC his H	CGC arg R
CTA leu L	CCA pro P	CAA gln Q	CGA arg R
CTG leu L	CCG pro P	CAG gln Q	CGG arg R

ATT ile I	ACT thr T	AAT asn N	AGT ser S
ATC ile I	ACC thr T	AAC asn N	AGC ser S
ATA ile I	ACA thr T	AAA lys K	AGA arg R
ATG met M	ACG thr T	AAG lys K	AGG arg R

GTT val V	GCT ala A	GAT asp D	GGT gly G
GTC val V	GCC ala A	GAC asp D	GGC gly G
GTA val V	GCA ala A	GAA glu E	GGA gly G
GTG val V	GCG ala A	GAG glu E	GGG gly G

PHYSICAL UNITS

°C	temperature in degrees Celsius
g	gram
m	metre
mol	mole
s	second
Ci	Curie [radioactivity; $3.7 \times 10^{10} \text{ s}^{-1}$ (disintegrations per second)]
Da	Dalton (relative molecular mass)
g	gravitational acceleration (9.81 ms^{-2})
hr(s)	hour(s)
l	litre (volume; 10^{-3} m^3)
M	molar concentration (mol.l^{-1})
min(s).	minute(s) (time)
S	Svedberg (sedimentation)
U	unit of enzymatic activity
V	volts
A	amperes

ORDER PREFIXES

d	deci	10^{-1}	k	kilo	10^3
c	centi	10^{-2}	M	mega	10^6
m	milli	10^{-3}	G	giga	10^9
μ	micro	10^{-6}	T	tera	10^{12}
n	nano	10^{-9}			
p	pico	10^{-12}			
f	femto	10^{-15}			
a	atto	10^{-18}			

CONTENTS

<u>INTRODUCTION</u>	1
CHAPTER 1: VACCINES AGAINST VIRAL DISEASES	2
1/ ADAPTIVE IMMUNE RESPONSES TO VIRAL VACCINES	2
1.1/ HUMORAL IMMUNE RESPONSES.....	2
1.2/ CELL-MEDIATED IMMUNE RESPONSES.....	3
2/ VIRAL VACCINES	4
2.1/ WHOLE VIRUS VACCINES.....	4
2.2/SUBUNIT VIRUS VACCINES.....	5
2.2.1/ <i>Purified proteins / synthetic peptides</i>	6
2.2.2/ <i>Recombinant viral vectors</i>	6
2.2.3/ <i>Genetic immunisation</i>	7
3/ SMAA COMPLEXES AS MULTIVALENT SUBUNIT VACCINES	9
 CHAPTER 2: THE PARAMYXOVIRIDAE: THE VIRUSES AND THEIR STRUCTURE	11
1/ THE VIRUSES	11
2/ PARAMYXOVIRUS STRUCTURE AND ENCODED PROTEINS	11
2.1/ VIRION STRUCTURE.....	11
2.2/ PARAMYXOVIRUS GENOME.....	13
2.3/ PARAMYXOVIRUS ENCODED PROTEINS.....	14
2.3.1/ <i>The nucleocapsid protein (NP/N)</i>	14
2.3.2/ <i>The phospho protein (P)</i>	14
2.3.3/ <i>The large protein (L)</i>	15
2.3.4/ <i>The matrix protein (M)</i>	16
2.3.5/ <i>The integral membrane proteins</i>	16
2.3.5.1/ <i>The attachment protein (HN/H/G)</i>	16
2.3.5.2/ <i>The fusion protein (F)</i>	21
2.3.5.3/ <i>The small hydrophobic protein (SH)</i>	24
 CHAPTER 3: HETEROLOGOUS GENE EXPRESSION	26
1/ INTRODUCTION	26
2/ MAMMALIAN-CELL(MC) EXPRESSION SYSTEMS	29

2.1/ CURRENT VECTORS FOR GENE TRANSFER INTO MCs	31
2.1.1/ VIRAL VECTOR SYSTEMS.....	31
2.1.1.1/ Papovavirus expression system.....	32
2.1.1.2/ Vaccinia virus expression system.....	33
2.1.1.3/ Adenovirus expression system.....	35
2.1.1.4/ Adeno-associated virus expression system.....	39
2.1.1.5/ Alphavirus expression system.....	40
2.1.2/ DIRECT DNA TRANSFER.....	42
2.1.2.1/ Choice of promoter and enhancer.....	43
2.1.2.2/ Intron present or absent & choice of 3' untranslated region (UTR).....	46
2.1.2.3/ Optimal translation initiation sequence.....	48
2.1.2.4/ Optimisation for mammalian codon usage.....	48
2.1.2.5/ Transient or stable expression of the protein.....	49
 CHAPTER 4: THE MAMMALIAN-CELL EXOCYTOTIC PATHWAY.....	53
1/ Introduction.....	53
2/ Targeting of proteins to the ER.....	53
2.1/ <i>Signal sequences</i>	53
2.2/ <i>SRP-dependent targeting</i>	55
3/ Translocation of polypeptides across the ER membrane.....	55
4/ Modification of newly translocated polypeptides within the ER.....	56
4.1/ <i>Polypeptide folding and oligomerisation</i>	56
4.2/ <i>Addition of carbohydrates to nascent polypeptide chains</i>	58
5/ Retention of proteins within the ER.....	59
6/ Transport of proteins from the ER to the <i>trans</i> Golgi network (TGN).....	60
7/ Transport of proteins from the TGN to lysosomes or the cell-surface & external milieu.....	61
 CHAPTER 5: OBJECTIVES OF THE PROJECT.....	62
 <u>MATERIALS & METHODS</u>	65
1/ Bacterial strains.....	65
2/ Mammalian tissue culture cells & antibodies.....	65
3/ Plasmids.....	66
4/ Agarose gel electrophoresis.....	66

5/ Extraction of DNA fragments from agarose gels.....	67
6/ DNA linkers and primers.....	67
7/ Polymerase Chain Reaction (PCR).....	69
8/ Processing of amplified cDNAs for ligation.....	69
9/ DNA ligations.....	69
10/ Preparation of competent bacteria.....	69
11/ Transformation of competent bacteria.....	70
12/ Preparation of plasmid DNA.....	70
13/ DNA sequencing.....	71
14/ Preparation of cationic liposomes for DNA transfection.....	71
15/ Staining protocol for the detection of β -galactosidase expressing cells.....	72
16/ Quantitative luciferase assay.....	72
17/ Production of vaccinia virus vTF7-3 stocks.....	72
18/ Titration of vaccinia virus stocks.....	73
19/ Transient protein production:	
(a) Expression from the T7 ϕ 10 promoter.....	74
(b) Expression from the hCMV IE promoter in 293 cells.....	74
(c) Expression from the hCMV IE promoter in Cos-7 cells.....	74
(d) Expression from the tetracycline-responsive promoter in 293 cells.....	75
20/ Synchronisation of mammalian cells at the G2/M phase.....	75
21/ Production of stable cell lines.....	75
22/ Determination of protein concentration.....	76
23/ SDS-PAGE and Western blot analysis.....	76
24/ Stripping of nitrocellulose or PVDF membranes.....	77
25/ Immunoprecipitation.....	77
26/ Immunofluorescence.....	77
27/ Colony hybridisation.....	78
28/ Chemical cross-linking reactions.....	79
29/ PNGase F treatment of whole cell extracts.....	79
30/ Tunicamycin treatment of transfected cells.....	79

RESULTS.....81

SECTION 1: ANALYSIS OF TRANSIENTLY EXPRESSED Pk-TAGGED SV5 GLYCOPROTEINS.....81

1.1/ CONSTRUCTION OF VECTORS FOR THE EXPRESSION OF Pk- TAGGED HN & F.....81

1.1.1 / Construction of the vectors encoding an N or C-terminal Pk-tag.....81

1.1.2/ Construction of the vectors encoding native SV5 HN and N or C-terminally Pk-tagged HN.....	85
1.1.3/ Construction of the vector for the expression of C-terminally Pk-tagged SV5 F.....	85
1.2/ TRANSIENT PROTEIN PRODUCTION FROM THE VECTORS WHICH ENCODE Pk-TAGGED HN & F.....	85
1.2.1/ Expression of Pk-tagged SV5 HN & F.....	85
1.2.2/ N-linked glycosylation analysis of Pk-tagged SV5 HN & F	86
1.2.3/ Localisation of Pk-tagged SV5 HN & F.....	87
1.2.4/ Oligomeric structure of Pk-tagged SV5 HN & F.....	88
1.2.5/ Antigenicity of Pk-tagged SV5 HN & F to mAbs raised against native SV5 HN & F, which recognise conformationally-dependent epitopes.....	89
1.3/ CONSTRUCTION OF VECTORS FOR THE EXPRESSION OF TRUNCATED Pk-TAGGED HN & F.....	90
1.3.1/ Construction of vectors for the expression of truncated Pk-tagged HN.....	91
1.3.2/ Construction of vectors for the expression of truncated Pk-tagged F.....	91
1.4/ TRANSIENT PROTEIN PRODUCTION FROM THE VECTORS WHICH ENCODE TRUNCATED Pk-TAGGED SV5 HN & F.....	92
1.4.1/ Expression and localisation of the truncated N & C-terminally Pk-tagged SV5 HN.....	92
1.4.2/ Expression and localisation of the truncated C-terminally Pk-tagged SV5 F.....	93
1.4.3/ N-linked glycosylation analysis of truncated Pk-tagged SV5 HN & F....	93
1.4.4/ Oligomeric structure of C-terminally Pk-tagged truncated SV5 F.....	94
1.5/ <u>SUMMARY</u>.....	94

SECTION 2: ATTEMPTS AT PRODUCING CELL-LINES WHICH EXPRESSED Pk-TAGGED SV5 HN.....95

2.1/ ATTEMPTS AT THE CONSTRUCTION OF CELL-LINES WHICH CONSTITUTIVELY EXPRESSED Pk-TAGGED SV5 HN.....	96
2.1.1/ Construction of EBV-based vectors which encode C - terminally tagged SV5 HN.....	96
2.1.2/ Transient protein production from the vectors pREP4HNctPk and	

pCEP4HNctPk.....	96
2.1.3/ Attempts at the production of stably expressing cell-lines using the vectors pREP4HNctPk and pCEP4HNctPk.....	97
2.1.4/ Construction of the EBV-based vector encoding C - terminally Pk-tagged SV5 HN with a reversed transcription unit.....	97
2.1.5/ Transient protein production from the vector pCEP4HNctPk (->->).....	98
2.2/ CONSTRUCTION OF TRANSIENT EXPRESSION VECTORS WHICH PRODUCE Pk-TAGGED SV5 HN VIA AN hCMV IE PROMOTER.....	98
2.2.1/ Construction of pcDNA-based vectors for the transient production of N & C-terminally Pk-tagged SV5 HN.....	99
2.2.2/ Production of N & C-terminally Pk-tagged SV5 HN from the pcDNA-based vectors.....	99
2.3/ ATTEMPTS AT THE CONSTRUCTION OF CELL-LINES WHICH INDUCIBLY EXPRESSED Pk-TAGGED SV5 HN.....	100
2.3.1/ Construction of pTET vectors which encode N & C - terminally Pk-tagged SV5 HN.....	100
2.3.2/ Transient protein production from the vectors pTETHNctPk and pTETHNntPk.....	101
2.3.3/ Construction of pTEP4 vectors which encode N & C - terminally Pk-tagged SV5 HN.....	101
2.3.4/ Transient protein production from the vectors pTEP4HNctPk (->->) and pTEP4HN-ntPk (->->).....	102
2.3.5/ Attempt at the establishment of stable cell-lines which inducibly express N-terminally Pk-tagged SV5 HN using the vectors pEBV-tTa/neo & pTEP4HNntPk (->->).....	102
2.3.6/ Construction of, and expression from, vectors which encode tTa/hygB.	103
2.3.7/ Construction of, and expression from, the EBV-based vectors which inducibly express the HNntPk glycoprotein and confer resistance to histidinol.....	105
2.3.8/ Construction of cell-lines which constitutively express tTa from an episomally maintained EBV-based vector.	105
2.3.9/ Attempt at the establishment of stable cell-lines which inducibly express N-terminally Pk-tagged SV5 HN, using the vectors pTEP4HNntPk (->->) or pTEP8HNntPk (-> ->) and a tTa expressing cell-line.....	106
2.4/ INFLUENCE OF THE DRIVER/RESPONDER VECTOR-TYPES ON	

TRANSIENT PROTEIN PRODUCTION FROM THE INDUCIBLE EXPRESSION SYSTEM.....	107
---	------------

2.5/ <u>SUMMARY</u>.....	109
---------------------------------	------------

SECTION 3: A NOVEL APPROACH FOR THE AMPLIFICATION OF THE COPY NUMBER OF EPISOMES STABLY MAINTAINED WITHIN PRIMATE CELL-LINES.....	112
--	------------

3.1/ CONSTRUCTION OF AN SV40 ORI-CONTAINING EBV-BASED VECTOR WHICH CONSTITUTIVELY EXPRESSES THE SV5 P PROTEIN.....	112
---	------------

3.2/ TRANSIENT PROTEIN PRODUCTION FROM THE VECTOR pCEP8/P (->->).....	113
--	------------

3.3/ MANIPULATIONS TO INCREASE THE EXPRESSION OF THE SV5 P PROTEIN ENCODED FROM EBV-BASED VECTORS.....	113
---	------------

3.3.1/ Influence of the orientation of the SV5 P gene transcription unit on the production of P protein in a transient expression analysis.....	113
--	------------

3.3.2/ Effect of synchronising the mammalian cells at the G2/M phase of the cell-cycle prior to transfection.	114
---	------------

3.4/ CONSTRUCTION OF THE pCT7-BASED VECTOR WHICH ENCODES THE SV40 LTA_g.....	115
---	------------

3.5/ TRANSIENT PRODUCTION OF THE SV40 LTA_g FROM THE VECTOR pCT7LTA_g.....	115
---	------------

3.6/ ESTABLISHMENT OF CELL-LINES WHICH EXPRESS THE SV5 P PROTEIN ENCODED FROM AN EBV-BASED VECTOR.....	116
---	------------

3.7/ TRANSIENT EXPRESSION OF SV40 LTA_g IN CELL-LINES WHICH EXPRESS THE SV5 P PROTEIN ENCODED FROM AN SV40 ORI-CONTAINING EBV-BASED VECTOR.....	116
--	------------

3.8/ <u>SUMMARY</u>.....	117
---------------------------------	------------

<u>DISCUSSION</u>.....	118
-------------------------------	------------

1/ Pk-tag as an epitope tag for the detection and purification of recombinant proteins.....	118
--	------------

1.1/ Analysis of transiently expressed Pk-tagged HN & F.....	118
---	------------

1.1.1/ Analysis of full length Pk-tagged HN & F.....	118
---	------------

1.1.2/ Analysis of truncated Pk-tagged HN & F.....	121
2/ Isolation of cell-lines which produce Pk-tagged HN.....	126
2.1/ Constitutive expression of Pk-tagged HN.....	126
2.2/ Inducible expression of Pk-tagged HN.....	128
2.3/ Influence of vector type on transient gene expression.....	130
3/ Amplification of EBV-based vectors.....	130
3.1/ Production of the SV5 P protein from an SV40 <i>ori</i> -containing EBV- based plasmid.....	132
3.2/ Production of the SV40 LTA _g protein.....	133
3.3/ Analysis of the transient amplification experiment.....	134
4/ Future developments.....	135
4.1/ Pk-tag and anti-Pk mAb for use as an epitope-tagging system.....	135
4.2/ Novel vectors for recombinant protein production in mammal.- cells..	136
<u>APPENDIX A</u>	139
<u>APPENDIX B</u>	141
<u>REFERENCES</u>	143

ABSTRACT

The work presented within this thesis expands upon the theme within this laboratory, of utilising epitope-labelled recombinant proteins for the construction of multivalent subunit vaccines. Mammalian-cell expression-vectors were constructed which encoded a 14 amino acid epitope-tag, termed Pk-tag. The genes encoding the haemagglutinin (HN) and fusion (F) glycoproteins (model type II and type I proteins respectively) from the paramyxovirus simian virus 5 (SV5), were inserted into the above vectors such that the sequence encoding the Pk-tag was present at the amino (N) or carboxy (C) terminus of SV5 HN, and the C-terminus of SV5 F. The genes were expressed in mammalian cells by utilising the vaccinia virus/T7 transient-expression system. Encouraging results were obtained which demonstrated that the addition of the Pk-tag to the N or C termini of SV5 HN, or to the C-terminus of SV5 F, did not prevent the production of; full length, N-linked glycosylated, oligomeric, natively folded and cell-surface localised Pk-tagged protein.

An attempt was made to produce secretable forms of Pk-tagged SV5 HN and F. For this purpose, a vector was constructed which encoded a truncated version of the SV5 F protein in which the C-terminal transmembrane anchor & cytoplasmic tail were deleted, but which still possessed the sequence encoding the C-terminal Pk-tag at the C-terminus of the ectodomain. Expression of this gene in mammalian cells resulted in the production of a protein which had undergone N-linked glycosylation and partial oligomerisation. No secretion of the truncated Pk-tagged protein into the external milieu was detected.

Furthermore, production of potentially secretable forms of N & C-terminally Pk-tagged SV5 HN was achieved by the construction of plasmid vectors in which the non-cleavable native HN signal sequence was replaced by a putative cleavable signal sequence from the Epstein Barr virus gp220/360 glycoprotein. Expression of the modified Pk-tagged HN genes in mammalian cells produced proteins of a lower mwt. than expected and which, apart from a small proportion of the N-terminally Pk-tagged molecules, did not possess N-linked oligosaccharides and were not recognised by conformationally sensitive mAbs. No secretion of the modified Pk-tagged HN into the external milieu was detected.

Following the first initial characterisation of the Pk-tagged HN & F, in which very encouraging results were obtained, an attempt was made to isolate cell-lines which constitutively or inducibly expressed Pk-tagged HN. Production of Pk-tagged HN could not be detected from constructs in which expression was driven from constitutive promoters. However, production of N-terminally Pk-tagged HN (but not C-terminally tagged HN) was detected when expression was driven by the tTa inducible expression system. As a further development to the tTa system, a 293 cell-line was isolated which expressed high-levels of functional tTa. The tTa-producing cell-line was subsequently utilised in an attempt to isolate cell-lines which inducibly produce N-terminally Pk-tagged HN. These cells are currently undergoing selection for drug resistance.

Further experiments were performed to try and develop a system whereby the low copy number of episomally-maintained EBV-based vectors present in a stable cell-line could be amplified to high copy number, whereby a subsequent increase in protein production would be envisaged. For this

purpose, SV40 ori-containing, EBV-based vectors were constructed in which the expression of the non-toxic SV5 P protein was under the control of the hCMV IE promoter/enhancer. Cell-lines were isolated which produced the SV5 P in various amounts. Transient amplification of the episomal copy number was attempted via a transient expression of the SV40 LTag, using plasmid DNA transfection. No subsequent increase in protein production was observed by way of a Western blot analysis.

INTRODUCTION

The long term aim of one of the projects within this laboratory has been concerned with the development of multivalent subunit vaccines, via what has been termed solid matrix-antibody-antigen or SMAA complexes. The work presented in this thesis, although continuing with the above theme, has been directed towards the production and analysis of epitope-labelled simian virus 5 (SV5) glycoproteins in mammalian cells. Subsequently this project has involved extensive recombinant DNA technology, recombinant protein production, and analysis of protein processing through the mammalian-cell exocytotic pathway. Therefore, the following introduction primarily concentrates upon the heterologous expression of proteins, particularly in mammalian cells, and upon the processing of glycoproteins derived from the paramyxovirus SV5.

Chapter 1 of the introduction briefly outlines present and possible future methods for the production of vaccines against viral disease and introduces the concept of producing multivalent subunit vaccines as SMAA complexes. Chapter 2 focuses on the paramyxoviruses, with emphasis being directed towards the paramyxovirus simian virus 5 (SV5) and its envelope glycoproteins: haemagglutinin-neuraminidase (HN) and fusion (F) proteins. One of the most important areas in recombinant protein production is choosing and developing a suitable expression system for the protein(s) of choice, chapter 3 outlines heterologous gene expression technology with particular emphasis being placed on vectors used for recombinant protein production from mammalian cells. Furthermore, a thorough understanding of the post-translational modifications which may occur to expressed proteins and how targeting of proteins can be manipulated within mammalian cells, is important to the development of some recombinant vaccines, and chapter 4 briefly presents some of the knowledge gained in the field of mammalian-cell protein exocytosis. Finally, in the last section of the introduction, the aims of the work undertaken in this thesis are stated.

CHAPTER 1

VACCINES AGAINST VIRAL DISEASE

1 / Adaptive immune responses to viral vaccines.

Protection from viral disease in vertebrates is due to the application of two defence mechanisms: non-adaptive & adaptive immune responses. Non-adaptive immune responses involves barriers to the primary infection, such as the external skin and mucous linings of the respiratory and gastrointestinal tract, and also circulatory phagocytic cells which may engulf and neutralise foreign bodies. Adaptive immune responses are composed of humoral and cell-mediated immune responses, primarily involving antibodies and cytotoxic T lymphocytes (CTLs) respectively.

1.1 / Humoral immune responses.

Humoral immune responses primarily defend a host organism against infection of its cells by invading viral pathogens (viral neutralisation). However, the humoral response may also target virus-infected cells for lysis by either the complement pathway or through a process termed antibody-dependent-cellular-cytotoxicity, or ADCC. The above responses are mediated through glycoprotein effector molecules termed immunoglobulins (Igs or antibodies, of which there are nine classes in man; IgA₁, IgA₂, IgD, IgE, IgG₁, IgG_{2a}, IgG_{2b}, IgG₃, IgG₄ & IgM), which are either membrane-bound or secreted from plasma cells (derived from activated B cells) and circulate throughout the body via the blood and lymph.

Antibodies, or to be more exact the antigen-binding site, or paratope, of the antibody, recognise small antigenic determinants termed epitopes which are present in the structure of

a foreign molecule (antigen). Such molecules tend to be proteins or modified proteins and, to a lesser extent, polysaccharides and smaller chemicals. Epitopes can either be continuous or discontinuous (Barlow *et al.*, 1986), whereby continuous epitopes consist of a linear sequence of amino acids (derived from the primary amino acid sequence), usually consisting of 5 to 10 amino acids in length (reviewed in Randall & Souberbielle, 1990). In contrast, discontinuous epitopes are derived from the juxtaposition of amino acids, distinct from each other in the primary amino acid sequence, mediated by the folding of the protein into its tertiary and quaternary structures. Discontinuous epitopes are therefore conformationally sensitive. The extent to which the humoral immune response protects against viral disease depends primarily upon the nature of the infecting viral agent.

1.2 / Cell-mediated immunity.

Due to the host-cell dependent nature of a virus replication cycle, viruses are, during their replication/assembly phase, concealed from the neutralising effects brought about by the humoral immune system. Therefore, a further defence against the spread of virus from infected cells, is through the specific targeting of such infected cells by T-cells. T-cells recognise infected cells through receptors (T-cell receptors, TCRs) present on their cell-surface, which interact with virus-derived oligopeptide fragments that are displayed on the surface of infected cells. This surface expression is mediated through an association of the viral-derived oligopeptides with cellular molecules termed major histocompatibility complexes, of which there are two types: MHC I & MHC II.

There are distinct subsets of T cells which are distinguished by the cluster differentiation marker (CD) they possess on their cell-surface. $CD4^- CD8^+$ T-cells interact with oligopeptides presented in association with MHC I molecules, whereas $CD4^+ CD8^-$ T-cells interact with oligopeptides associated with MHC II molecules. Other surface molecules are also involved in the interaction of T-cells with their target cells which are reviewed in Hanke & Randall, 1994.

Primarily $CD8^+$ cells are involved in the lysis of infected cells and are thus termed cytotoxic T-lymphocytes (CTLs), whereas $CD4^+$ cells are primarily involved in immune

modulation through the release of soluble effector molecules termed cytokines, which regulate processes such as the activation of CTLs and the proliferation of B-cells into antibody-secreting plasma cells, although these functions are not mutually exclusive (reviewed in Hanke & Randall, 1994).

2 / Viral vaccines

Although the spread of many viral diseases has been successfully controlled by improvements in public health facilities, active immunisation, through the administration of vaccines, has also played a major role in curtailing the dissemination of viral disease, a classic example of which is the elimination of variola virus from the general population, and recently the elimination of poliovirus from North America. A further need for effective viral vaccines is the poor chemotherapy which is currently available for viral diseases.

A vaccine is "material originating from a micro-organism that induces an immunologically mediated resistance to disease" and not necessarily to infection (Mims, 1988). Examples of vaccine types currently in use, and also of current experimental vaccines are depicted in figure 1. An ideal vaccine should possess the following characteristics: induce resistance to disease, be capable of oral administration, induce a life-long immunity to disease after only one dose, be safe, highly stable, cheap to produce and give combined protection against multiple diseases. However, there exists, at present, no such ideal vaccine and current vaccines possess only a combination of the above features.

2.1 / Whole virus vaccines

Conventional methods for viral vaccine preparation has involved the use of whole virus particles. As depicted in figure 1, vaccines derived from whole virus particles can be divided into two categories: inactivated particles or infectious particles. Inactivated virus vaccines are usually prepared by treatment of wildtype virus with protein cross-linking agents such as formalin and β -propiolactone. Infectious virus vaccines can be separated into 3 categories; attenuated virus, non-pathogenic related virus, and wild type pathogen administered by a route which doesn't result in disease (examples of inactivated and

Whole Virus Vaccines

Inactivated virus - *Poliovirus (Salk)*

- *Influenza virus*
- *Rabies virus*
- *Hepatitis A virus*

Infectious virus -> **Attenuated**

- *Measles virus*
- *Mumps virus*
- *Rubella virus*
- *Poliovirus (Sabin)*
- *Yellow fever virus*

-> **Virulent**

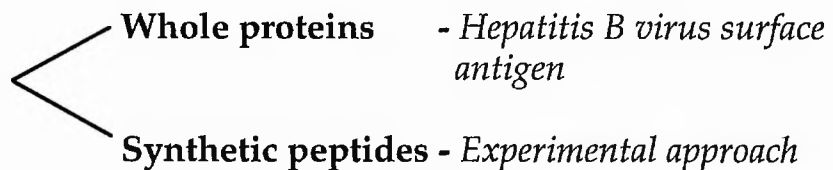
- *Adenovirus type 5*

-> **Related**

- *Vaccinia virus*
- *Calf Rotavirus*

Subunit Virus Vaccines

Amino acid-based Administration



Nucleic acid-based Administration (DNA or RNA)

Whole proteins/
Polyepitopes - Experimental approach

Recombinant Viral Vectors

Whole proteins/
Polyepitopes - Experimental. Vectors include; *Vaccinia virus*, *Adenovirus*, *Herpes simplex virus*, *Retroviruses* and *Alphaviruses*

Figure 1 - Commonly administered types of viral vaccine for human use, with examples of viral pathogens, and some recent novel approaches to future vaccine development. See text for details.

infectious viral vaccines in use to date are shown in figure 1). Attenuated viruses have traditionally been produced by serial passage of wild type virus in a non-host-related cell-line, with subsequent screening for potential attenuated variants. However, molecular biology techniques now allow the possibility of selective mutagenesis for certain viruses. The above vaccine preparation approaches have obviously met with great success but they both have their advantages and disadvantages, as compared below.

Inactivated vaccines - produced from virulent particles.

- generally multiple doses required.
- large number of particles required.
- generally induce a poor CTL immune response (see above).
- possibility of incomplete inactivation.
- generally can be administered to immunosuppressed persons.
- preparations are usually very stable.
- spread to unvaccinated persons has not been demonstrated.
- possibility of enhancement of disease upon natural infection.
- the inactivation process can lead to destruction of neutralising epitopes.

Infectious viral vaccines - possibility of reversion to virulent phenotype.

- possibility of spread to unvaccinated persons.
- risk of developing persistent or latent infections with some vaccines.
- generally can not be administered to immunosuppressed people.
- possibility of contamination of cell-lines by virulent microbes.
- tend to be less stable than inactivated vaccines.
- usually stimulates both humoral & CTL responses (see above).
- small number of particles required.
- generally fewer doses are required for adequate protection.
- possibility of enhancement of disease upon natural infection.

2.2 / Subunit virus vaccines

Subunit vaccines, depicted in figure 1, are a relatively new approach to vaccine development and can be administered either as; peptides/purified proteins, by a novel nucleic acid-based approach termed genetic immunisation, or by infectious recombinant

virus vectors. Successful use of subunit vaccine technology has been demonstrated with the current hepatitis B virus vaccine which is produced from recombinantly produced surface antigen expressed in yeast (Valenzuela *et al.*, 1982).

2.2.1/ Purified proteins / synthetic peptides.

Advantages of using purified proteins/synthetic peptides for immunisation include that only preselected immunogens are administered and that the vaccine is non-infectious. In the latter case, this is an important safety consideration if vaccines are being developed against agents such as retroviruses. However, there are numerous disadvantages which include; (1) it can be expensive and technologically challenging to produce large amounts of purified proteins, (2) it may be difficult to determine which proteins/peptides are inducing protective immune responses, (3) MHC polymorphism within the human population has to be taken into account in the design of peptide-based vaccines and finally (4) the administration of proteins/peptides on their own is generally only weakly-immunogenic. However, there are numerous methods for presenting proteins/peptides to the immune system and for enhancing their immunogenicity (a process termed adjuvanticity), which have been experimentally demonstrated to induce good humoral and importantly, in some approaches, CTL responses. These include; aluminium gels (the only adjuvant currently licensed for human use), quil A, muramyl dipeptide (derived from *M. tuberculosis*), solid-matrix-antibody-antigen complexes (Randall & Young 1991, and discussed below) and liposomes (reviewed in Gupta *et al.*, 1993).

2.2.2/ Recombinant viral vectors.

Included in the category of subunit virus vaccines are recombinant viruses which have been altered using recombinant DNA techniques to express proteins, or portions thereof, from unrelated viruses and are thus acting as infectious carriers for the administration of subunit vaccines. The advantages of such a system include; (1) due to infection by the virus vector, genes are expressed within the host cell and the proteins are therefore properly processed (e.g folding and glycosylation) and will therefore tend to

induce appropriate humoral and CTL immune responses, and (2) vectors such as adenovirus type 5 and vaccinia virus are, or have been, routinely used for mass immunisation programmes and their safety, although not absolute, is relatively good. However, there are additional problems to those already described above for peptide/protein-administered subunit virus vaccines, which include; (1) the size of the viral genome can restrict the size of the insert genes, and (2) protective immune responses will probably be mounted to the virus vector as well as to the subunit protein, thus rendering repeated administrations unlikely (reviewed in Randall & Souberbielle, 1990). [More detail on the construction of recombinant viruses and their use in heterologous protein expression is given in chapter 3, section 2.1.1].

2.2.3/ Genetic immunisation.

One of the most recent, novel and exciting approaches to vaccine advancement has been the development of technology for direct-gene-transfer into living tissue, primarily as part of a DNA plasmid molecule. Therefore, this section will be more expansive in detailing some of the current knowledge in this field. This approach of direct-gene-transfer has been referred to as genetic immunisation or nucleic acid immunisation. The advantages of such an approach include;

- 1) Since viral proteins are produced within the host cell, the proteins will acquire host-cell post-translational modifications and will therefore adopt the native conformation.
- 2) Intracellular production of the desired proteins will also therefore result in the elicitation of CTL responses as well as humoral responses. Indeed, humoral and/or CTL responses have clearly been demonstrated against the rabies virus G protein (Xiang *et al.* 1995), influenza virus HA & NP proteins (Ulmer *et al.*, 1993 and Fynan *et al.*, 1993), HIV-1 gp120 (Fuller & Haynes, 1995) and the hepatitis B surface antigen (Davis *et al.*, 1995 and Mancini *et al.*, 1996).
- 3) It is very cheap, safe and straight-forward to produce large amounts of plasmid DNA to very high purity.

- 4) DNA can be stored lyophilised at room temperature and is stable for at least one year in this condition, thereby allowing easy storage and transport.
- 5) Plasmid DNA is relatively easy to manipulate and therefore desirable characteristics can be added to gene products or, conversely, deleterious features removed.
- 6) Immunisation can be multivalent simply by producing a mixture of plasmids encoding different immunogens, or by the production of poly-CTL-epitope proteins which can be constructed from known CTL-epitopes of various viruses or tumour antigens (Thomson *et al.*, 1996).

However, as with the vaccination strategies mentioned above, there are foreseeable problems which could arise from using plasmid DNA as a means of genetic immunisation. Such problems include;

- 1) The introduction of DNA into the cell nucleus may result in the insertion of the plasmid, or portions thereof, into the host cell genome thereby possibly eliciting insertional mutations in essential genes, or upregulating potential oncogenes. However, integration of DNA into genomes generally requires the active replication of the genome which does not commonly occur in post-mitotic cells such as muscle cells, a common site for injection of DNA. Indeed, it has been demonstrated that plasmid DNA injected into mouse muscle cells remains in an extrachromosomal form (sensitive PCR and Southern blotting techniques could not detect any integrated plasmid DNA sequence), and has not undergone replication as elucidated by the presence of the bacterial methylation pattern (reviewed in Danko & Wolff, 1994).
- 2) Anti-DNA immune responses may be generated to transfected plasmid molecules resulting in auto-immune diseases such as systemic lupus erythematosus (SLE). However, it is apparently very difficult to illicit a humoral immune response against intact DNA and one study involving DNA immunisation could not detect the presence of any anti-DNA antibodies (reviewed in Robertson, 1994 and Xiang *et al.* 1995).
- 3) Depending upon the tissue which takes-up and expresses the foreign gene, expression may not be prolonged enough to develop a protective immune response. Indeed it has been demonstrated that direct injection of plasmid DNA into exposed brain, liver, spleen, uterus, stomach, lung or kidney did not result in any detectable gene expression, but then direct

injection of such tissues is not a practicable situation for human vaccination. However, in a perhaps more relevant scenario it has been demonstrated that muscle cells which have taken-up circular plasmid DNA, via direct injection, can express the gene product (in this case luciferase) for at least 2 years, or for at least for 4 months if the plasmid was linearised prior to injection (Wolff *et al.* 1990 and reviewed in Danko & Wolff, 1994).

4) However, theoretical problems may arise due to prolonged expression such as; the induction of tolerance, auto-immunity, anaphylaxis, hyperimmunity or autoaggression (reviewed in Robertson, 1994). Circumstances where such phenomenon may occur due to prolonged expression of the antigen may be over come by selective targeting of the DNA to tissues which rapidly degrade the DNA or perhaps by incorporating genes encoding slow-acting toxic proteins into the plasmid vectors thereby eliminating transfected cells after a short period of expression. Another possible approach to achieving either short-term or prolonged gene expression is by controlling such expression via the use of inducible promoters. Perhaps, to date, the most likely candidate for such an inducible promoter is the tetracycline-responsive promoter developed by Gossen *et al.* (1995) whereby gene expression is induced by the addition of tetracycline derivatives (more detail on this system is given in chapter 3, section 2.2.2.1).

5) Uptake of naked DNA into tissue is an inefficient process and in the case of direct injection into muscle cells, necrotizing agents such as snake toxins or local anaesthetics (such as bupivacaine) have been administered to facilitate the uptake of DNA. However, a non-invasive approach using DNA-coated gold particles fired onto the epidermis of the skin via a "gene gun" has shown promising results (Fynan *et al.*, 1993).

3 / SMAA complexes as multivalent subunit vaccines.

Solid-matrix-antibody-antigen (SMAA) complexes have been suggested to be a potential carrier system for the presentation of single or multiple microbial subunits to the immune system (Randall, 1989). The basis of the design, which is depicted in figure 2, consists of a solid matrix to which are bound appropriate monoclonal antibodies (mAbs) to the antigen of choice. The solid matrix, to date, has consisted of a killed & fixed suspension

of the Cowan A strain of *Staphylococcus aureus* (Randall & Young, 1988), or the adjuvant alum (Randall *et al.*, 1993b). In the case of *S. aureus* the mAbs interact with the solid matrix via binding of the Ig Fc region to protein A present on the surface of the *S.aureus*. Purified antigen can then be incorporated into the complex via an interaction with the specific monoclonal antibodies, thereby also facilitating another purification step upon the antigens of choice.

SMAA complexes have been demonstrated to induce both vigorous humoral and class I-restricted CTL immune responses (Randall & Young, 1988), where immunisation of mice with SMAA complexes containing internal & envelope proteins from the paramyxovirus simian virus 5 (SV5) protected mice from subsequent infection by SV5 (Randall *et al.*, 1988). Furthermore, clearance of a persistent SV5 infection in mice was achieved through the action of CD8⁺ effector T-cells, induced by the administration of SMAA complexes containing internal proteins and envelope glycoproteins from SV5 (Randall & Young, 1991).

In a recent development to the above SMAA complex technology, Hanke *et al.* (1992, 1994a, 1994b and 1995) have demonstrated that a single mAb, which recognises a 14 amino acid oligopeptide epitope (derived from the P/V proteins of SV5 (Southern *et al.*, 1991)), can be used in the construction of SMAA complexes containing multiple epitope-labelled-antigens (figure 2, panel C). The mAb has been referred to as anti-Pk and the epitope to which it interacts has been labelled Pk-tag. The advantages of such a system include; (1) antigens for which there are no monoclonal antibodies available can be Pk-tagged and incorporated into SMAA complexes, and (2) large scale production of only one monoclonal antibody is required..

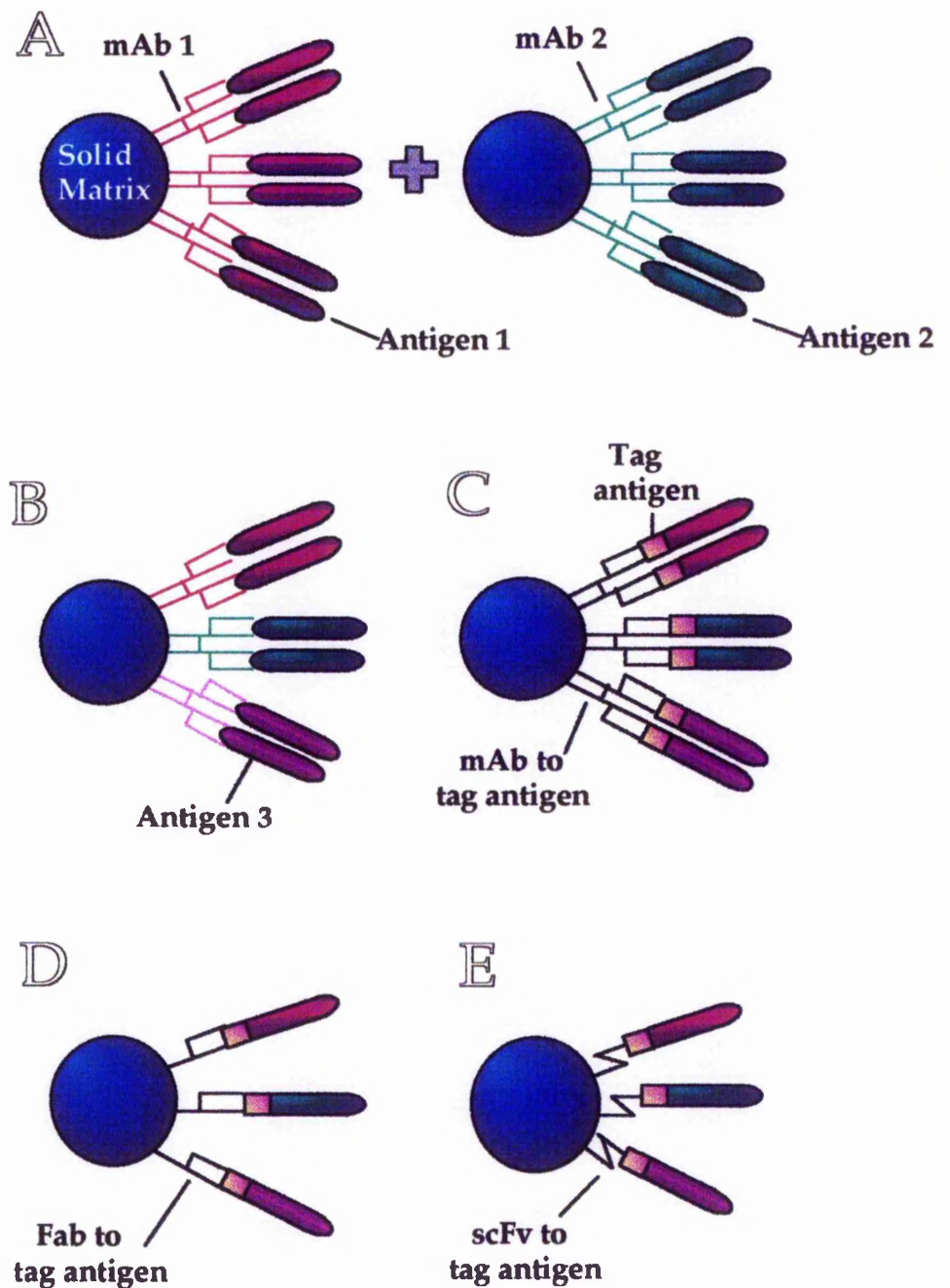


Figure 2: Schematic of possible SMAA complexes which could be utilised as multivalent vaccines.

A - Different mAbs are attached to individual solid matrices prior to the binding of the respective antigens. The resulting SMAA complexes may be pooled together before being employed as potential immunogens. **B** - The solid matrix is saturated with a mixture of mAbs prior to the addition of the respective antigens. **C** - A single tag-specific mAb and tag-linked antigens are utilised for the construction of SMAA complexes. **D & E** - The same as **C**, except that Fab or single chain Fv fragments of antibodies, respectively, are employed. Furthermore, additional mAbs to host cell antigens, and/or immunostimulating protein such as cholera toxin B subunit, may be included to enhance particular types of immune response (not shown). Modified from Hanke (1993).

CHAPTER 2

THE PARAMYXOVIRIDAE: THE VIRUSES AND THEIR STRUCTURE.

1 / The Viruses.

The viruses contained within the family Paramyxoviridae possess a cell-derived envelope and contain a linear negative strand RNA genome. Virus members within the family Paramyxoviridae therefore belong to the same group of viruses comprising the families; Rhabdoviridae, Arenaviridae, Bunyaviridae and Orthomyxoviridae. The Paramyxoviridae and the Rhabdoviridae differ from the other 3 families in that their genome is non-segmented.

In 1993 the family Paramyxoviridae was divided by the International Committee on the Taxonomy of Viruses into two subfamilies (figure 1): namely the Paramyxovirinae and the Pneumovirinae. The subfamily Paramyxovirinae contains three genera, the Parainfluenzavirus, Rubulavirus and the Morbillivirus, whilst the subfamily Pneumovirinae possesses one genera, the Pneumovirus. This new classification was based upon several criteria; genome organisation, biological activities of the encoded proteins, sequence relationship of the encoded proteins and morphological characteristics.

2 / Paramyxovirus structure and encoded proteins.

2.1 / Virion Structure (figure 4).

Paramyxoviridae are pleiomorphic viruses which, although generally observed to be spherical in nature, can form filamentous particles. The viruses contain a host-cell-derived

lipid membrane into which are inserted membrane-bound virus-encoded glycoproteins which extend approximately 8 to 12 nm from the envelope surface (Lamb & Kolakofsky, 1996). The envelope surrounds a "scaffold" comprised of a virus-encoded Matrix (M) protein, inside of which is located the helical ribonucleoprotein (RNP) core.

THE FAMILY PARAMYXOVIRIDAE

THE SUBFAMILY PARAMYXOVIRINAE

GENUS *Parainfluenzavirus* - Sendai Virus (SeV).
 - Human parainfluenza virus type 1 & 3 (hPIV1/3).
 - Bovine parainfluenza virus type 3 (bPIV3).

GENUS *Rubulavirus* - Simian virus 5 (SV5).
 - Mumps virus (MuV).
 - Newcastle disease virus (NDV).
 - Human parainfluenza virus types; 2,4a & 4b.
 (hPIV2/4a/4b).

GENUS *Morbillivirus* - Measles virus (MeV).
 - Dolphin morbillivirus (DMV).
 - Canine distemper virus (CDV).
 - Phocine distemper virus (PDV).
 - Rinderpest virus (RiV).

THE SUBFAMILY PNEUMOVIRINAE

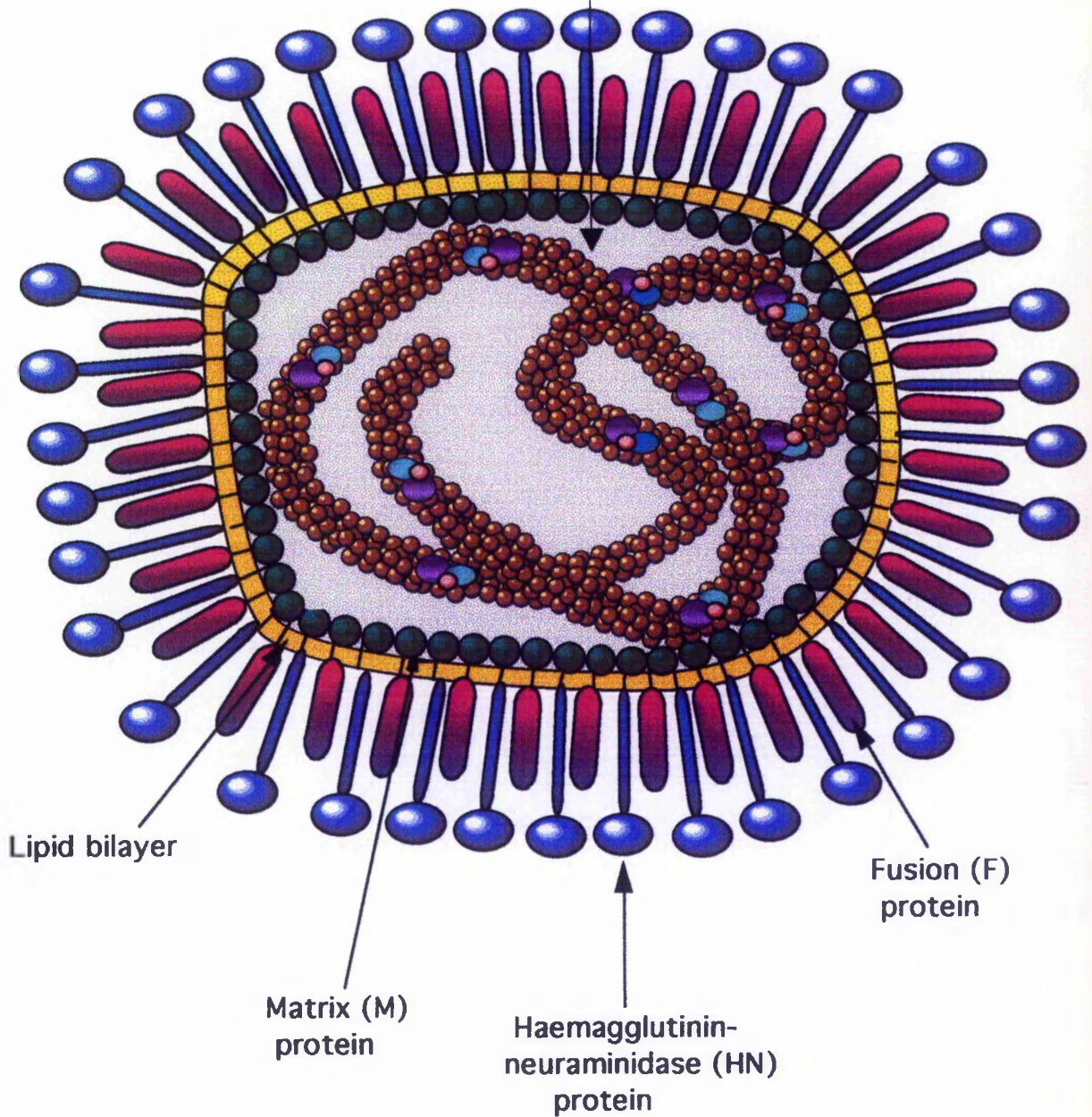
GENUS *Pneumovirus* - Human respiratory syncytial virus (hReSV).
 - Bovine respiratory syncytial virus (bReSV).
 - Pneumonia virus of mice (PVM).
 - Turkey rhinotracheitis virus (TRV).

Figure 3- Organisation of the family Paramyxoviridae, showing subfamilies, genera and representative examples of viruses, abbreviations shown as used throughout this thesis. Adapted from Lamb & Kolakofsky 1996.

Figure 4: Diagrammatic representation of the structure of the simian virus 5 (SV5) virion (adapted from Young, 1991).

SV5 consists of a single strand RNA genome of negative polarity to which are associated the nucleocapsid proteins (NP). Also associated with the NPs are the phospho (P) proteins and the large (L) proteins which form the viral polymerase complex. The V protein has been shown to be associated with the nucleocapsid but definitive evidence for this association has not been published. Surrounding this nucleocapsid is a matrix produced from the homologous association of the matrix (M) protein which interacts with the host-cell-derived lipid bilayer and to the bilayer-inserted virally encoded glycoproteins - Haemagglutinin-Neuraminidase (HN) and Fusion (F). Not shown is the membrane-bound small hydrophobic (SH) protein, whose function is unknown.

Nucleocapsid:
[Nucleoprotein (NP-●), associated with Large (L-●),
(V-●) and Phospho-(P-●) proteins, complexed with
negative sense RNA genome]



The ribonucleoprotein core (sometimes referred to as the nucleocapsid core) consists of the single-stranded RNA genome of negative polarity (approximately 15Kbp), which is tightly associated with the nucleoprotein (NP) to which is adhered the phosphoprotein (P) and Large (L) proteins. This RNP core, rather than the free genomic RNA, is the template for all RNA synthesis (Hamaguchi *et al.*, 1983) performed by the polymerase complex which is comprised of, at least, the P and L proteins.

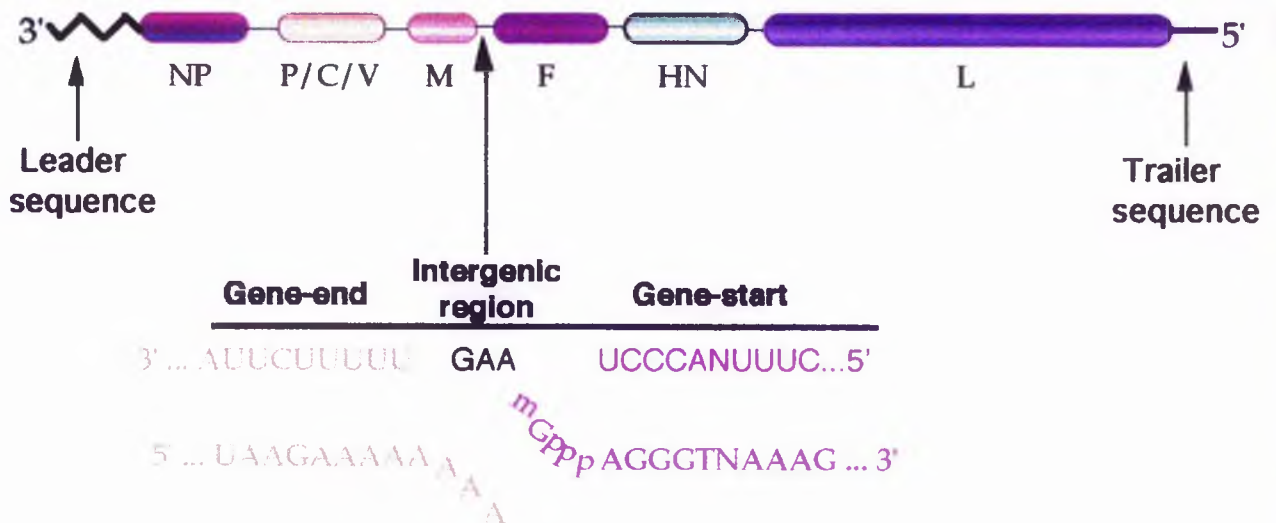
2.2 / Paramyxovirus genome (figure 5).

The negative sense linear genomic RNA of the Paramyxoviridae has to serve two functions within the replicative cycle of the viruses: first to act as a template for synthesis of the positive stranded antigenome, which is an intermediate in the replicative cycle, and second to act as a template for the synthesis of messenger RNAs (mRNAs). As previously mentioned, the genome of the Paramyxoviruses is a nonsegmented single-stranded RNA of negative polarity and approximately 15Kbp in length. The genome codes for between 6 and 10 polypeptides which are transcribed, by a virally encoded polymerase, with decreasing frequency as the distance from the 3' leader sequence increases. The complete sequence of the RNA is known for many members of the Paramyxoviridae, including; Sendai virus, Human parainfluenza virus 3, Simian virus 5, Mumps virus, Measles virus, Canine distemper virus and Respiratory syncytial virus (Gallinski *et al.*, 1991). Using this sequence information, sequences which may comprise coding and non-coding regions of the genome can be predicted and possible functions assigned. The genome has been shown to possess 3' (leader sequence) and 5' (trailer or negative leader sequence) extra-cistronic regions of approximately 50 nucleotides, which appear to be control regions for regulating transcription and translation. Further potential control elements have been located at the start and end of individual Paramyxovirus genes which are semi-conserved between different viruses (Gallinski, 1991). Moreover, regions putatively involved in transcription/replication control are the non-transcribed intercistronic regions (ICR), which again are semi-conserved between the Paramyxoviridae and in the case of Measles virus,

Figure 5: Genome organisation of representative members from each genus of the family Paramyxoviridae.

For Sendai virus, regions putatively involved in transcription/replication control are indicated in bold lettering, with genomic sense and mRNA species illustrated. Gene size is approximately drawn to scale (adapted from Lamb & Kolakofsky 1996).

Parainfluenzavirus - Sendai virus



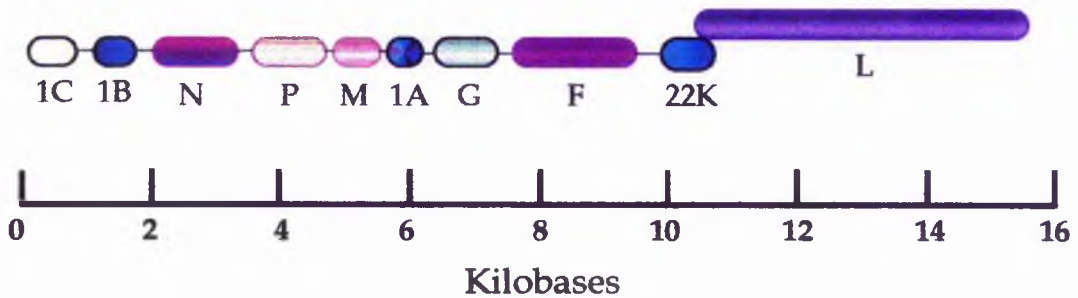
Rubulavirus - Simian virus 5



Morbillivirus - Measles virus



Pneumovirus - Respiratory syncytial virus



Sendai virus and PIV3 consists of a trinucleotide sequence. The ICR of the other Paramyxoviruses vary from between 1-52 nucleotides.

2.3 / Paramyxovirus encoded proteins.

2.3.1 / The nucleocapsid Protein (NP).

The nucleocapsid proteins of the Paramyxoviridae are acidic in nature, range in molecular weight between 53 and 58Kda (Lamb & Kolakofsky, 1996) and are the most abundant proteins to be found within the RNP (Lamb *et al.*, 1976). NP has several functions in virus replication including; genome encapsidation (thus rendering the genome RNase-resistance), an association with the P-L polymerase complex during transcription/replication, and a possible interaction with the matrix protein during virion assembly (Markwell and Fox, 1980). During viral replication, it is hypothesised that the polymerase complex recruits NP and that along with the elongating genome, is incorporated into a growing RNP. There does not appear to be a unique interaction with genomic RNA as NP expressed on it's own in cell-lines appears to form nucleocapsid-like structures, utilising cellular RNA, of equal density to viral nucleocapsids (Blumberg *et al.*, 1983, Buchholz *et al.*, 1993). The N-terminal three quarters of Sendai virus NP appears to be all that is necessary for RNP assembly (Buchholz *et al.*, 1993) and indeed is well conserved among related viruses. This region is composed of 2 separate domains, the N-terminal domain is positively charged and presumably interacts with the genomic RNA, with the other region tending to be hydrophobic in nature, and thus may be involved in the NP:NP interactions (Morgan E.M *et al.*, 1984). The C-terminal region of NP has been found to be vital in the replication of an encapsidated cDNA-encoded Sendai virus genome where nucleocapsids prepared from C-terminally deleted NP could not act as templates for new rounds of genome replication (Curran *et al.*, 1993).

2.3.2 / The Phosphoprotein (P).

As the name suggests, the P proteins are highly phosphorylated. The P gene of the Pneumovirinae and hPIV1 encode a single polypeptide, whilst in the case of some members

2nd copy - white

have please

—



8 9 10



X



3 2 1



R 1 abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMN OPQRSTUVWXYZ ABCDEFGHIJKLMN OPQRSTUVWXYZ £123456789

R 2 abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMN OPQRSTUVWXYZ ABCDEFGHIJKLMN OPQRSTUVWXYZ £123456789

R 3 abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMN OPQRSTUVWXYZ ABCDEFGHIJKLMN OPQRSTUVWXYZ £123456789

R 4 abcdefghijklmnopqrstuvwxyz ABCDEFGHIJK
LMNOPQRSTUVWXYZ £1234567890

R 5 abcdefghijklmnopqrstuvwxyz&
ABCDEFGHIJKLMN OPQRSTU
WXYZ £1234567890



1A

10 _____

9 _____

8 _____

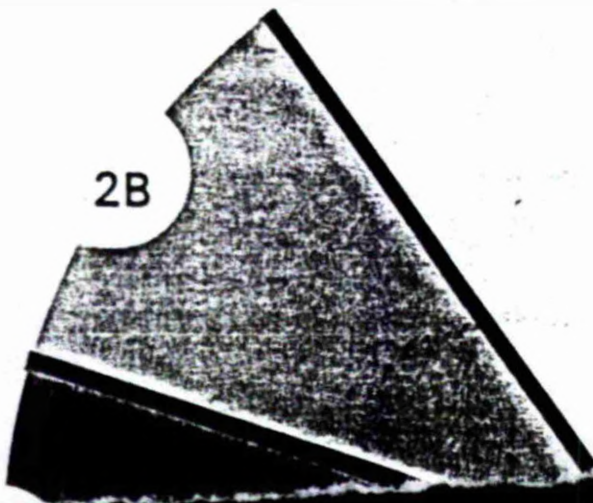
7 _____

6 _____

5 _____

T

2B



of the Paramyxovirinae multiple polypeptides are produced. This is due to overlapping reading frames from internal start codons (C proteins produced) and RNA editing where pseudo-templated addition of nucleotides can occur, which result in frame-shifting during translation and a separate species of mRNA produced (which encodes V proteins, reviewed in Lamb & Kolakofsky, 1996).

The length of the P proteins vary considerable within the Paramyxoviridae, ranging from 241 amino acids in the *Pneumoviruses* to 603 amino acids in the *Parainfluenzaviruses* and *Morbilliviruses*. The P protein is a regulatory protein which has a key role in transcription and replication. In unison with the viral L protein it forms the viral polymerase complex (P-L). Moreover, in association with soluble NP, the complex P-NP is formed which presumably maintains the NP in a soluble state until the interaction with the nascent RNA (Horikami *et al.*, 1992 and Parks, 1994, Precious *et al.*, 1995). Furthermore, P may play a major role in the interaction of the L protein with the template RNA by binding to the nucleocapsid NP whilst also maintaining an association with the L protein (Ryan and Portner, 1990).

The exact function of the aforementioned C and V proteins is in dispute but they may play a role in the fine tuning of the transcription / replication modes of at least some members of the Paramyxoviridae (Curran *et al.*, 1992 & 1994 and Precious *et al.*, 1995).

2.3.3 / The Large (L) protein.

Many L genes from the Paramyxoviridae family have been sequenced and shown to express a large polypeptide of approximately 2200 amino acids. The large size and low abundance (approximately 50 copies per virion) of the L protein in the virion suggested it's role as the RNA-dependant RNA polymerase (RDRP) (reviewed in Lamb & Kolakofsky, 1996). Overall there appears to be very little sequence homology except in five short regions of high homology, which is also conserved between other RDRPs, for example of Vesticular stomatitis virus (Poch *et al.*, 1990). To date, no structure-function studies have been carried out for paramyxovirus L proteins.

2.3.4 / The Matrix (M) protein.

The M protein is the most abundant protein in the virion and appears to play a major role in the budding of newly produced viral particles. The molecular weights range from between 38Kda to 42Kda for the different members of the family, with the protein possessing a basic and slightly hydrophobic nature, albeit lacking any known membrane-spanning domain (reviewed in Lamb & Kolakofsky, 1996). Fractionation studies and electron micrographs indicate that the M protein is located underlying the lipid membrane as a peripheral membrane protein, and can self-associate to form paracrystalline sheets and tubes *in vitro*. An association with the RNP has also been postulated (the basic nature of M may thus be important in the interaction with the acidic NPs of the RNP, reviewed in Lamb & Kolakofsky, 1996). Furthermore, it appears that the M protein may interact with the cytoplasmic tails of the virally-encoded integral membrane glycoproteins (Sanderson *et al.*, 1993) and evidence of an interaction with actin has been found (Bohn *et al.*, 1986). For Sendai virus, recently synthesised M appears to be highly phosphorylated, yet M protein located within matured virions is unphosphorylated. A clear role for this phosphorylation phenomenon has not yet been established (Lamb & Choppin, 1977).

2.3.5 / The integral membrane proteins.

The pneumoviruses and some members of the rubulaviruses (SV5 and MuV) encode three polypeptides which are located within the envelope of virions and infected cells; the attachment protein, fusion protein and small hydrophobic protein. Other members of the Paramyxoviridae family only encode the attachment and fusion proteins. As their names imply, these envelope proteins are involved in the adsorption, penetration and intracellular spread of the viruses. For the most part in this thesis, greater detail will be given for the envelope proteins derived from SV5.

2.3.5.1 / The attachment proteins.

The attachment proteins of the Paramyxoviridae are primarily involved in the adsorption of virions to target cells. For the *Parainfluenzaviruses* and *Rubulaviruses*, the

attachment protein (designated Haemagglutinin-Neuraminidase, HN) has both haemagglutinating activity, for the binding of virions to sialic acid containing receptors, and neuraminidase activity (which cleaves sialic acid and possesses an acidic pH optima of between pH 4.8 & 5.5), presumably for the release of progeny virions from infected cells, by prevention of self-aggregation (reviewed in Lamb & Kolakosky, 1996). In contrast, the *Morbilliviruses* have an attachment protein (designated Haemagglutinin, H) which possesses only haemagglutinating capabilities, but in contrast to HN, is not believed to bind to sialic-acid-containing receptors due to the viruses restricted host range, and the lack of a neuraminidase activity. Indeed, the specific cellular receptor for measles virus has recently been elucidated as being CD46 (Naniche *et al.*, 1993). The *Pneumovirus* RS virus attachment protein (designated, G) has neither haemagglutinating or neuraminidase activities (Matthews, 1982; Gruber and Levine, 1983) and the cellular receptor has not yet been defined. A further function of some Paramyxoviridae attachment proteins is to aid in the membrane fusion and penetration stage of virion entry into cells. These functional differences observed between members of the Paramyxoviridae are reflected in the overall poor amino acid sequence homology between the proteins.

3D Structure - In contrast to the influenza virus HA and NA glycoproteins (Wilson *et al.*, 1981 and Varghese *et al.*, 1983), none of the Paramyxoviridae envelope glycoproteins have been, to date, successfully crystallised and X-ray analysed, although soluble forms of some of the HN glycoproteins have been produced which may well turn out to be more easily crystallised (Parks & Lamb 1990). However, the attachment proteins have been extensively analysed using biochemical and immunological methods. Indeed, tryptic digest analysis and 2D electron-microscopy has shown that the HN molecule is composed of a trypsin sensitive stalk region and a trypsin resistant head region which is composed of four subunits, and is analogous to the structure defined for the influenza virus NA. Sequence alignments between Influenza NA and the Paramyxovirus attachment proteins also reveal that amino acids involved in the neuraminidase activity of NA are conserved within the HN proteins (Colman *et al.*, 1993). The majority of these amino acids are located in 4 conserved

regions, 3 of which are absent from the MuV H protein, thereby supporting the role of these regions in the neuraminidase activity of HN.

Current evidence seems to suggest that the tetrameric structure for the attachment protein of some members of the Paramyxoviridae, (including; SV5, SeV, MuV and NDV) is composed of a mixture of covalently and non-covalently linked dimers (Markwell & Fox, 1980 and Ng *et al.*, 1989). The proportion of the dimers which covalently form tetramers or merely non-covalently associate to form tetramers may, in part, be due to expression level and hence the concentration of the dimers found within the ER. Low concentration of HN, as found early in infection produced predominantly non-covalently associated tetramers and conversely, high concentration, as found late in infection produced predominantly covalently bound tetramers (Morrison *et al.*, 1990). Stable formation of these tetramers is apparently dependant upon a functional trans-membrane domain, cytoplasmic region, and a correctly folded ectodomain of the HN monomer (Ng *et al.*, 1990; Parks & Lamb, 1990 and Parks & Lamb, 1991).

Monomeric subunit - In SV5, the HN monomer is produced as a polypeptide of 565 amino acids and has an unmodified predicted mwt. of ~ 62Kda (Hiebert *et al.*, 1985). The protein is generally divided into three regions (figure 6) comprised of the; 17 amino acid cytoplasmic tail, 19 amino acid trans-membrane domain (which co-functions as a signal sequence) and the 529 amino acid ectodomain. Furthermore, for the majority of the Paramyxoviridae, including; SV5, NDV, MeV, SeV, ReSV and hPIV3, the molecule has been shown to be orientated within the virion/cellular membrane in a type II topology (reviewed by Morrison, 1988 and illustrated in figure 7), analogous to the Influenza virus NA.

The attachment molecules are glycosylated by the addition of N-linked oligosaccharides (figure 6) and also by O-linked oligosaccharides in the case of the Pneumovirus G protein (Gruber & Levine, 1985 and Wertz *et al.*, 1985). In SV5 and MeV, oligosaccharides are added to all of the 4 potential N-linked glycosylation sites (Ng *et al.*, 1990 and Hu *et al.*, 1994) which gives the protein an apparent molecular weight of approximately 68-75Kda, depending on the cell-type in which expression is occurring

Figure 6: Diagrammatic representation of simian virus 5 (SV5) HN & F glycoproteins (adapted from Lamb & Kolakofsky, 1996).

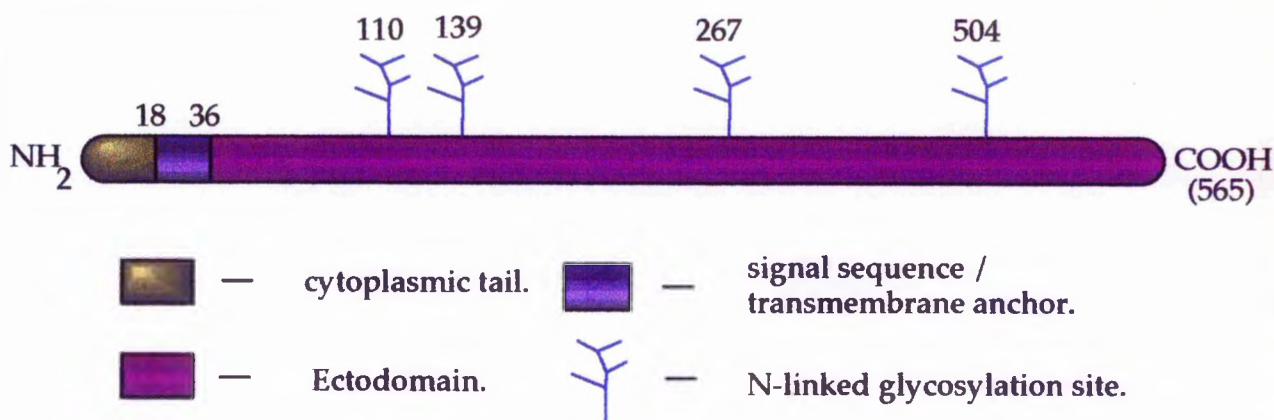
Panel A) *Schematic of the Haemagglutinin-Neuraminidase (HN) glycoprotein from SV5.*

SV5 HN possesses an N-terminal cytoplasmic tail of 17 amino acids followed by a 19 amino acid hydrophobic signal/stop-transfer sequence which targets and anchors the protein into the ER membrane, in a type II topology (see also figure 7). The ectodomain of the protein contains four potential sequons for the addition of N-linked oligosaccharides, all of which are utilised.

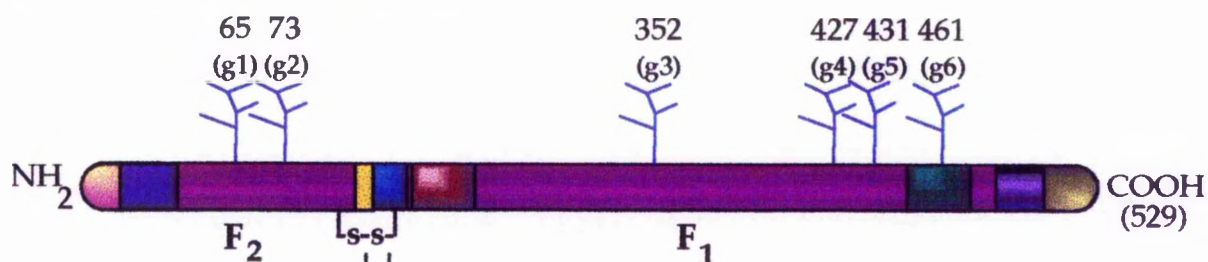
Panel B) *Schematic of the Fusion (F) glycoprotein from SV5.*

SV5 F possesses a 19 amino acid N-terminal signal sequence and membrane anchor which directs, and initially anchors, the nascent polypeptide to the ER. At the C-terminal region of the molecule are depicted the 19 amino acid cytoplasmic tail and the 38 amino acid membrane-anchor. Cleavage of the signal sequence occurs resulting in a type I membrane topology (see also figure 7). The ectodomain contains six potential sequons for the addition of N-linked oligosaccharides, all of which are utilised. Also depicted are the heptad repeat regions and the cleavage activation region along with representative examples of the fusion peptide amino acid sequence of other Paramyxovirinae, depicting the close homology of the peptide between the different members.

SV5 - HN (type II glycoprotein)



SV5 - F (type I glycoprotein)



																						% ID
SV5	-	F	A	G	V	V	I	G	L	A	A	L	G	V	A	T	A	A	Q	V	T	100
HPIV2	-	F	A	G	V	V	I	G	L	A	A	L	G	V	A	T	A	A	Q	I	T	90
MUMPS	-	F	A	G	I	A	I	G	I	A	A	L	G	V	A	T	A	A	Q	V	T	85
MEASLES	-	F	A	G	V	V	L	A	G	A	A	L	G	V	A	T	A	A	Q	I	T	80
R' PEST	-	F	A	G	V	A	L	A	G	A	A	L	G	V	A	T	A	A	Q	I	T	80
CDV	-	F	A	G	V	V	L	A	G	A	A	L	G	V	A	T	A	A	Q	I	T	75
HPIV1	-	F	F	G	A	V	I	G	T	I	A	L	G	V	A	T	A	A	Q	I	T	75
HPIV3	-	F	F	G	G	V	I	G	T	I	A	L	G	V	A	T	S	A	Q	I	T	70
NDV	-	F	I	G	A	I	I	G	S	V	A	L	G	V	A	T	A	A	Q	I	T	70
SENDAI	-	F	F	G	A	V	I	G	T	I	A	L	G	V	A	T	S	A	Q	I	T	70
		1			5						10				15					20		

Figure 7: Schematic depicting the membrane insertion of type I & II glycoproteins (adapted from Shaw *et al.*, 1988).

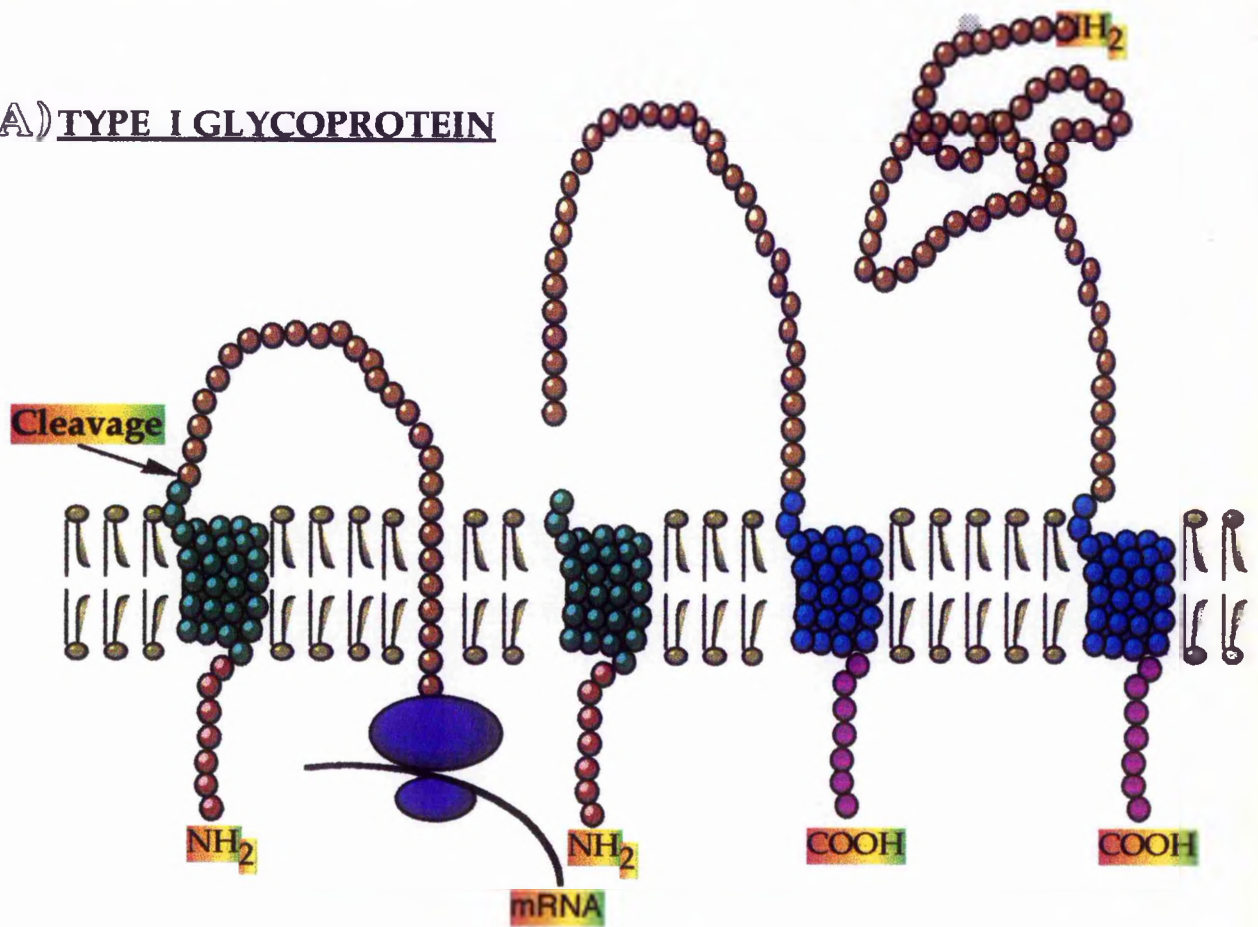
Panel A) *Type I glycoprotein.*

A type I glycoprotein possesses an N-terminal signal sequence of approximately 10-20 amino acids, which directs the ribosomes to the ER. A membrane-spanning stop-transfer sequence component of the signal sequence, anchors the polypeptide into the ER membrane followed by the continual translocation of the growing polypeptide chain into the ER lumen. A stop-transfer sequence present at the C-terminus halts further translocation of the polypeptide into the ER lumen, and translation typically terminates after the addition of a short C-terminus cytoplasmic tail. Cleavage after the first stop-transfer sequence, by signal peptidase, releases the N-terminus into the ER lumen.

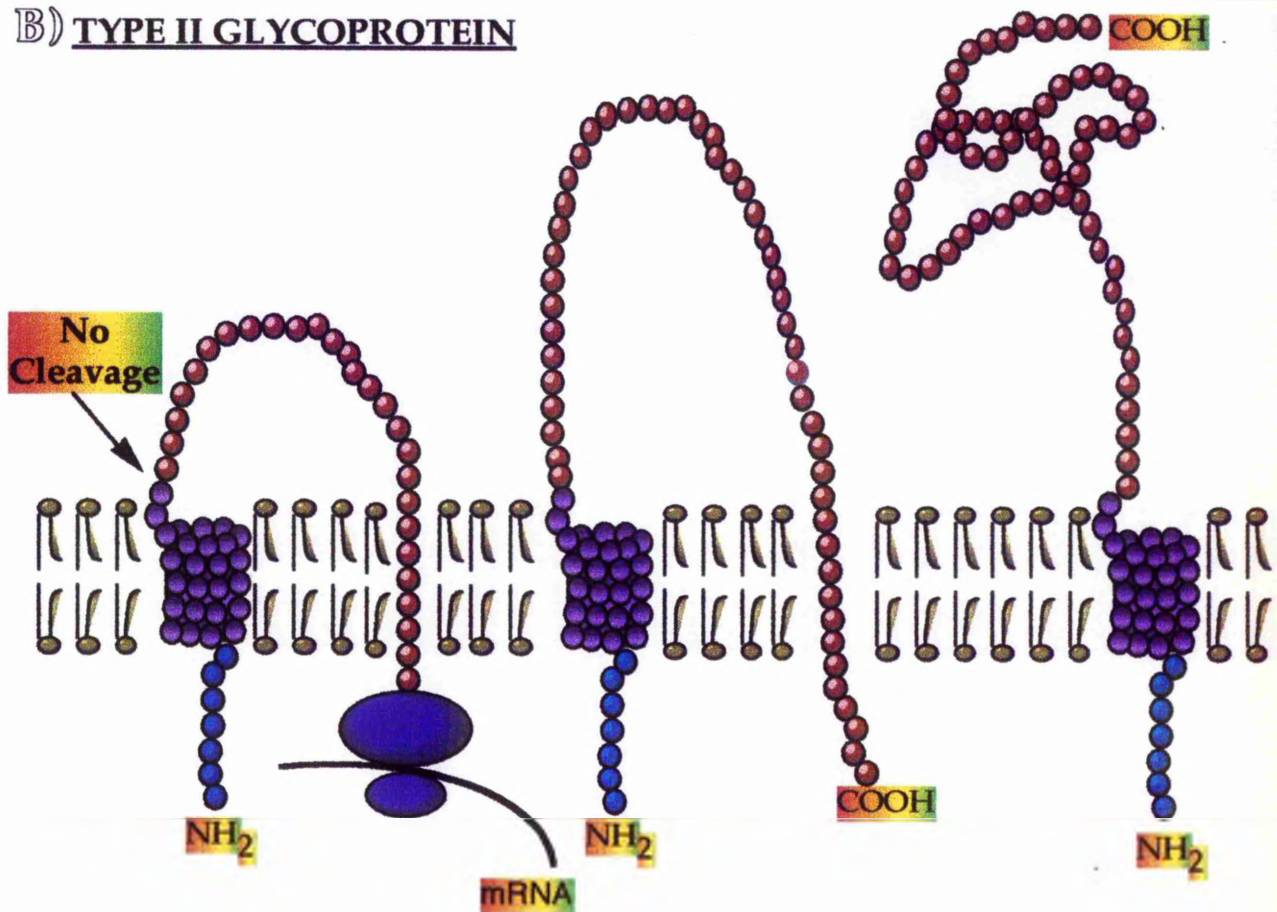
Panel B) *Type II glycoprotein.*

A type II glycoprotein possesses a short N-terminus cytoplasmic tail which is followed by a single hydrophobic region which serves as both a signal sequence, to direct the ribosomes to the ER, and as a membrane anchor. No cleavage after the membrane anchor occurs, and translocation of the nascent polypeptide chain into the ER lumen continues until translation is terminated.

A) TYPE I GLYCOPROTEIN



B) TYPE II GLYCOPROTEIN



(Paterson *et al.*, 1985). The ReSV G protein is 298 amino acids long with an unglycosylated mwt. of ~33Kda and a glycosylated mwt. of ~84Kda to 90Kda, half of which is attributable to O-linked glycosylation (Collins, 1991). For the HN and H proteins of SeV and MuV it has been shown that glycosylation is required for incorporation into the virions (Nakamura *et al.*, 1982 and Herrler & Compans, 1983). However, for NDV, glycosylation apparently wasn't required for virion incorporation, but was required for infectivity (Morrison *et al.*, 1981). Site-directed mutagenesis of the N-linked glycosylated sites in the SV5 protein showed varying effects on protein folding, transport and oligomerisation depending on the site of the mutation and the number of mutations introduced into the one molecule. Removal of all the glycosylation sites, or the expression of WT protein in the presence of the N-linked glycosylation inhibitor, tunicamycin, resulted in the production of a protein which; was not recognised by conformationally-sensitive mAbs, did not oligomerise to form the native structure, was not transported to the cell-surface, and which accumulated in the ER due to a stable association with the resident ER protein GRP78-BiP. The removal of the third (g3) glycosylation site at amino acid 267 had the most deleterious effect on protein folding. This appears to be in contrast to a similar study carried out by Hu *et al.* (1994) in which it was observed that mutations of the individual N-linked sites in the MeV H protein had little effect on folding and oligomerisation. However, in both cases, it was still apparent that native folding was due to a synergistic effect from the individual carbohydrate chains.

Glycine residues (important in forming bends and α -helices) and cysteine residues (important in intra-and-intermolecular bonding) within the attachment proteins have also been implicated as being important in the folding and oligomerisation process of protein maturation, and are well conserved within the Paramyxoviridae (Morrison & Portner, 1991 and Hu & Norrby, 1994). Furthermore, it has been demonstrated that the HN protein of SV5 undergoes fatty acid acylation via the addition of myristic and palmitic acid residues, possibly to the cysteine residue within the cytoplasmic tail (Veit *et al.*, 1989). The exact function of these modifications is unclear, but they may possibly contribute to the stability of the protein within the membrane (Morrison & Portner, 1991).

Intracellular Processing (greater detail on protein trafficking can be found in chapter 3) - Using pulse-labelling experiments and sedimentation studies Ng *et al.* (1989) have demonstrated that SV5 HN monomers oligomerise, in the ER, with a half-life ($T_{1/2}$) of 25-30 mins., prior to transport to the golgi network. This is significantly slower than the oligomerisation rates determined for Influenza virus HA & VSV G which have $T_{1/2}$ of 7-10 mins. and 6-8 mins. respectively (Gething *et al.*, 1986 and Doms *et al.*, 1987). The transport rate of the HN to the medial golgi network was demonstrated using EndoH analysis in which EndoH resistant forms of HN were detected after approximately 1hr, but an accurate determination could not be made due to the presence of multiple species of HN after EndoH treatment (Ng *et al.*, 1989). Using pulse-labelling experiments and immunofluorescence analysis it has been shown, for the SV5 HN, that the protein is turned over within the cell with a $T_{1/2}$ of ~2hrs, and is lost from the cell surface with a $T_{1/2}$ of ~ 50 mins (Ng *et al.*, 1990) which is similar to that found for SeV and MuV (Roux *et al.*, 1985 and Waxham & Wolinsky, 1986). This turn-over appears not to be dependent on the presence of other SV5 proteins, as HN expressed in CV1 cells infected by an SV40-recombinant also showed internalisation of HN (data not shown-Ng *et al.*, 1989). Furthermore, this internalisation was antibody independent. However, this loss from the cell-surface is not observed for hPIV3 HN which apparently is stably expressed at the cell-surface (Leser *et al.*, 1996).

Using gold-labelling electronmicroscopy immunocytochemistry experiments, Leser *et al.* have further shown that SV5 HN is internalised via clathrin-coated pits and that the internalisation appears to be of the non-recycling type, thereby resulting in the degradation of HN (included in this category are the Tumour Necrosis Factor, Interleukin 2 and Epidermal Growth Factor receptors, Leser *et al.*, 1996 and references therein). Furthermore, Leser *et al.* (1996) demonstrated that the SV5 HN protein was found to co-localise with transferrin (a marker for early endosomes), and bovine serum albumin (a marker for late endosomes). Prior to this study, Ng *et al.* (1989) showed that HN also co-localised with ovalbumin, a marker for lysosomes, thereby demonstrating the most likely degradative pathway for internalised SV5 HN. As the cytoplasmic tail of SV5 HN does not possess a tyrosine residue which has been implicated as being the critical residue for internalisation,

as present in the cytoplasmic tails of many proteins, the exact signal for internalisation of HN still remains to be elucidated, although a study by Ector *et al* (1992) demonstrated that the trans-membrane domain and the first six amino acids C-terminal to the trans-membrane domain were critical for internalisation. A possible role for the involvement of other SV5 proteins, as stated above can be ruled out, but an involvement with cellular proteins in directly inducing internalisation can not be eliminated.

Using co-precipitation studies of radio-labelled cells, Ng *et al.* (1989) showed that monomeric SV5 HN possessed a specific and transient ($T_{1/2}$ of 20-25 mins.) association with the resident ER chaperone protein GRP78-BiP. However, malformed monomers were found to be stably associated with GRP78-BiP ($T_{1/2}$ of >6hrs), thereby preventing the passage of malformed proteins to the golgi network.

Role in the Fusion process - An initial demonstration of the possible role of the Paramyxoviridae attachment protein in the fusion process was illustrated by Portner *et al.* (1987) where mAbs to the SeV HN were produced which inhibited virion fusion but did not affect the haemagglutinating or neuraminidase activities of the HN protein. Apart from SV5, all the other members of the Paramyxoviridae require the presence of the attachment protein for viral fusion (reviewed in Lamb & Kolakofsky, 1996 and references therein). From chimera studies, Deng *et al.* (1994) have shown that for hPIV2 & NDV HN proteins, the trans-membrane domain and the first 82 residues of the ectodomain are important for binding to F and aiding the fusion process. Furthermore, experiments carried out by Bousse *et al.* (1994) have demonstrated that a single point-mutation (asn->lys) in the proposed globular head region of hPIV1 HN was the source of the difference between a high and low fusogenic strain of hPIV1. No differences in the primary amino acid sequence of the F proteins were demonstrated.

2.3.5.2 / The Fusion Proteins.

The fusion (F) proteins are the major proteins of the Paramyxoviridae involved in cell-cell and virus-cell fusion, which occurs in a pH-independent manner (Choppin &

Compans, 1975). This is in contrast to the pH-dependant (pH5.0-5.5) fusion carried out by the Influenza virus HA (White *et al.*, 1983). With the exception of SV5, all members of the Paramyxoviridae require the co-expression of the homotypic HN/H protein for efficient fusion to take place (reviewed in Lamb & Kolakofsky, 1996). The requirement for co-expression of homotypic HN/H with the F protein may be due to the binding of the HN/H protein to its receptor molecule, where-upon a conformational change in the associated F protein occurs, resulting in a subsequent exposure of the fusion peptide. For SV5, contact of F with the cellular membrane or an unknown receptor may be the triggering factor.

Analysis of cloned mRNAs reveal that the F monomer ranges in length from between 529-580 amino acids (reviewed in Lamb & Kolakofsky, 1996). There is some degree of homology in the primary amino acid sequence throughout the family, especially for glycine, proline and cysteine residues, which is translated to an observed conservation of structural features (Morrison, 1988 and Morrison & Portner, 1991). The monomer possesses an N-terminal cleavable signal-sequence, and a C-terminal trans-membrane domain, which thus confers a type 1 membrane topology to the full length molecule (figures 6 & 7). The F-protein possesses between 3 and 6 potential N-linked glycosylation sites, with all six sites being occupied in the case of the SV5 F protein (Bagai & Lamb, 1995). Mutation of individual sites in the F₂ region of SV5 F resulted in a negligible change in surface expression, however, mutation of sites in the F₁ region resulted in minor to major changes in intracellular transport and protein stability (Bagai & Lamb, 1995). Furthermore, Collins & Mottet, (1991) have demonstrated that unglycosylated ReSV F undergoes inefficient cleavage. There is no evidence for O-linked glycosylation being involved, however, at least for NDV F, fatty acid acylation may be involved in the post-translational modification of the F-protein, but a function for this has not been elucidated. However, mutation of the cysteine residue in the cytoplasmic domain of the influenza virus HA protein (also a type I glycoprotein), which prevents palmitoylation of the protein, did not prevent surface expression, virion assembly or infectivity although, except for one mutant, all the viruses were attenuated in a mouse model (Jin *et al.*, 1996) .

The F-protein is synthesised as an inactive precursor (F_0) which is activated by proteolytic cleavage and trimming. This activation process is carried out by host-cell proteases in the trans-golgi network to form the disulphide linked subunits F_1 & F_2 (Scheid & Choppin, 1974). From cross-linking experiments, these monomers then go on to form non-covalently linked oligomers, possibly homotrimers (Russell *et al.*, 1994). From the nature of the cleavage site (shown in figure 6) and the site of cleavage within the cell, furin, a subtilisin-like protease, has been implicated as the enzyme responsible for the initial cleavage event (with the exception of Sendai virus). This initial cleavage is then followed by a trimming of the F_2 subunit by a carboxypeptidase (reviewed in Lamb & Kolakofsky, 1996).

The cleavage-activation process is a prerequisite for biological activity of the F-protein. Anderson *et al.* (1992) have shown that for the ReSV F, a mutant which does not undergo cleavage; was glycosylated (but not to complex forms), did not form oligomers, had a long association with the ER resident protein GRP78-BiP, was not recognised by conformation-specific mAbs, and had very little surface expression. Cleavage of F_0 into F_1 & F_2 exposes a hydrophobic domain, termed the fusion peptide (or fusion related ectodomain-FRED), which is 20 amino acids in length and is located at the N-terminus of the F_1 subunit. The fusion peptides from members of the Paramyxovirinae share between 70 and 90% amino acid homology (Paterson & Lamb, 1987 and depicted in figure 6). Evidence for this region being able to insert into membranes has been demonstrated using hydrophobic photo-affinity labelling probes (Novick & Hoekstra, 1988), and by the ability of the SV5 fusion peptide to act as a membrane anchor (Paterson & Lamb, 1987; Paterson & Lamb, 1990 and Parks & Lamb, 1990). Intriguingly, Horvath & Lamb (1992) have demonstrated that the invariant nature of the amino acids within the fusion peptide may have been selected so as to reach a compromise between fusion activity and viral replication, since mutation of glycines 3,7 & 12 to alanines dramatically increased syncytial formation.

Also located on the monomers are 2 regions of α -helix motifs, termed heptad repeats (figure 6), located immediately adjacent to the fusion peptide and the transmembrane anchor domain. A study by Sergel-Germano *et al.* (1994) demonstrated that these two

regions may be important in protecting the fusion peptide from the aqueous environment, and have perhaps a more direct role in the fusion process.

As has been examined for the HN/H proteins, the integrity of the cytoplasmic tail in the intracellular transport and biological activity of F has been analysed. A study by Sergel & Morrison (1995) using deletion mutants of the NDV F demonstrated that removal of the C-terminal tail produced a protein which underwent minimal proteolytic cleavage and resulted in a lack of syncytial formation. Deletion of the C-terminal half of the tail, resulted in a protein which was proteolytically cleaved but which did not result in syncytial formation. Removal of the C-terminal quarter, brought back 30% of syncytial forming activity, with WT characteristics been observed by a mutant deleted of the C-terminal eighth.

As has been previously mentioned, all members of the Paramyxoviridae, with the exception of SV5, require co-expression of homotypic HN/H & F to produce efficient fusion activity. An examination of the regions on MeV F which are involved with the interaction with its homotypic HN/H has been carried out by Wild *et al.* (1994). Using chimerae studies between CDV F and MeV F they observed that a region of 45 amino acids (Cys 337 -> Arg 381) located within the cysteine-rich region (figure 6) was critical for the interaction. Interestingly, SeV F disulphide links have been recently elucidated by Iwata *et al.* (1994) where 2 loops are produced. If this can be shown to extrapolate to MeV F then these loop regions are to be found within this critical region of binding to H, and therefore perhaps the loops interlock with similar "pocket" regions on the homotypic attachment protein.

2.3.5.3 / The Small Hydrophobic Protein (SH).

The *Rubulaviruses*, MuV & SV5, and members of the *Pneumoviruses* encode a third integral membrane protein, called the SH protein (or 1A in the case of ReSV), figure 5 (reviewed in Lamb & Kolakofsky, 1996). Detectable expression of an SH protein has been demonstrated for all the above viruses except in the case of MuV. The predicted amino acid sequence of SV5 SH indicates a protein of 44 amino acids in length, comprised of three domains: an N-terminal hydrophilic sequence, an intermediate hydrophobic sequence

which is of sufficient length to span a membrane, and a short (5 amino acids) hydrophilic C-terminal region. This arrangement would confer a type II topology on the SH protein (Hiebert *et al.*, 1988). *Pneumovirus* SH (or 1A) contains 64 amino acids and possesses similar domains and topology to the above SV5 SH. In contrast, ReSV SH is apparently modified by the addition of N-linked glycosylation on the C-terminal ectodomain, which adds a further 7-23Kda in mwt compared with the unglycosylated protein (reviewed in Lamb & Kolakofsky, 1996). A function for SH has not yet been elucidated, but it is found to co-localise with HN in SV5 infected cells, but is excluded from the virion (Hiebert *et al.*, 1988).

CHAPTER 3

HETEROLOGOUS GENE EXPRESSION

1 / Introduction.

The realisation that individual gene products could be produced within a heterologous host occurred with the advent of recombinant DNA technology in the early 1970s. At that early time point, the idea that genes could be inserted into plasmids, mobilised into bacteria with protein subsequently being produced was indeed revolutionary, with much of modern-day biotechnology and molecular biology being in its debt. However, since those early pioneering years, many different expression systems have been created for the heterologous expression, or over-expression, of introduced genes. Examples of such expression systems and some of their advantages & disadvantages are shown in table 1.

An important stage in any project where heterologous expression of a protein is required is the choice of host cell. Moreover, the choice of host cell is largely dependant on several factors, including; the nature of the protein, the purpose of protein production, the cost involved and the facilities & expertise available. The former two factors are considered to be the most important. The nature, or property, of the protein involved in the study can generally be grouped into 2 areas (table 2); nature of the polypeptide, and modifications/3D structure of the polypeptide as found in the native host cell. Generally, natively-unmodified polypeptides 80 to 500 amino acids in length which contain few cysteine residues and are not highly hydrophobic in nature should be well expressed in a bacterial system, especially if expressed as a fusion protein and/or are secreted into the external milieu (Goeddel, 1990).

However, large, highly modified, hydrophobic, mammalian cell-surface glycoproteins are unlikely to be expressed well in bacteria and an alternative system, such as insect cells or mammalian cells, would probably be more appropriate. However, the reason for protein production , as discussed below, will also influence the choice of host cell for expression.

Host Cell	Example	Advantages	Disadvantages
Bacterial	<i>E. coli</i> <i>B. subtilis</i>	Rapid growth, cheap, genetics well known, good for secretion, large amounts of protein made. Quick to set-up the system.	Few post-translational modifications. Poor codon useage for mammalian genes. Proteolytic degradation. Protein may possess N-terminal methionine/formyl methionine
Yeast	<i>S. cerevisiae</i> <i>P. pastoris</i>	As for bacterial, plus some post-translational modifications. Useful for protein-protein interaction studies.	Only a basic post-translational modification system present.
Fungi	<i>Aspergillus</i>	Rapid growth, cheap, some post-translational modifications.	As for yeast.
Insect	Sf9	Cheaper than mammalian cells. Mammalian-like post-translational modifications. Generally more protein produced than mammalian cells.	More expensive than yeast or bacterial systems. Genetics less well known. More time consuming than yeast or bacterial systems.
Protozoan	<i>Leishmania</i>	Rapid growth, cheap, some post-translational modification. Large amounts of protein produced.	Genetics less well known. Only basic post-translational modifications occur.
Mammalian	Hela, Cos-7, 293, CHO, BalbC, BHK.	Many post-translational modifications. Can observe phenotypic changes & protein-protein interactions.	Expensive. Time consuming to produce stable cell-lines. Slow growth rate. Generally poor yields of protein.

TABLE 1 - Examples of host cells available for the heterologous expression of recombinant proteins, listing some of their potential advantages and disadvantages.

NATURE OF POLYPEPTIDE	MODIFICATIONS/3D STRUCTURE
Small peptides, < 80 amino acids in length.	Phosphorylation.
Secreted polypeptides, 80-500 amino acids in length.	Glycosylation.
Secreted & surface polypeptides, > 500 amino acids in length.	Fatty acid acylation.
non-secreted polypeptides, > 80 amino acids in length.	Disulphide bond formation.
hydrophobicity/hydrophilicity of the polypeptide.	Oligomerisation.
	Specific proteolytic cleavage.

TABLE 2 - Factors related to polypeptide structure and post-translational modification which influence the choice of heterologous expression system for protein production. Note - not all of the possible post-translational modifications are listed.

Most applications of recombinant proteins fall into three broad categories

- i) Structural studies.
- ii) Functional studies.
- iii) Biotechnological development, e.g vaccines, hormones & pharmaceuticals.

Again the choice of host-cell depends on the nature of the protein. Not all structural studies on natively expressed mammalian proteins have to be produced from proteins expressed from a mammalian-cell system. Many natively unmodified intracellular mammalian proteins, or parts thereof, have been expressed in bacteria in large amounts and successfully crystallised and X-ray analysed (Ghosh *et al.*, 1995). However, most mammalian and viral surface glycoproteins are extensively post-translationally modified. Where these modifications confer, in part, the 3D structure of the molecule, an expression system will have to be chosen which carries out these natively-found modifications. The above statement

is also true for proteins which are being produced for functional studies, in that higher eukaryote proteins which are not extensively modified can be functionally produced in bacteria. Likewise, highly complex, integral membrane oligomeric enzymes found in mammalian cells, may only be successfully expressed, in a functionally active form, in mammalian cells or other higher eukaryotes.

Biotechnological use of expression systems for the production of recombinant proteins depends upon the proteins' use. For example, for some purposes antibodies can be readily produced against highly complex mammalian proteins which are produced in a malformed form from a bacterial expression system, either as a full length polypeptide, or portions thereof. However, antibodies which are produced, for example, to neutralise viral infectivity, may have to be raised against a natively folded protein which must, therefore, be expressed in a suitable expression system.

In summary, the choice of host-cell depends upon the nature of the protein and the intended use of the protein. For a number of applications for recombinant proteins, the generalisation "make host-like proteins in host-like systems" can be applied.

The rest of this chapter will detail current knowledge in the field of mammalian-cell expression. A general review on some of the other expression systems currently available, as detailed in table 1 is given in Goeddel, (1990) (bacteria, and yeast), Zhang *et al.*, (1995) (*Leishmania*), Gellissen *et al.*, (1992) & Cregg *et al.*, (1993) (*Pichia pastoris* expression system), and Summers & Smith, (1987) & O'Reilly, (1992) (Insect cells).

2 / Mammalian-cell expression systems.

Due to the aforementioned problems, considerable effort has been made in constructing systems for the heterologous expression of proteins in mammalian cells. The diversity of mammalian-cell expression technology has dramatically increased over the past 10-15 years such that now, for many purposes, expression in mammalian-cells is routine and a first choice system rather than been viewed as a tedious last-choice option.

There are numerous reasons for expressing cloned eukaryotic genes in mammalian-cells, these include (Sambrook *et al.*, 1989);

- 1) To verify the identity of a cloned gene by using immunological or functional assays to detect the encoded protein.
- 2) To generate large amounts of proteins that are normally only available in insufficient quantities from natural sources, e.g hormones, antibodies.
- 3) To examine the synthesis and intracellular transport of proteins when expressed in various cell-types.
- 4) To express genes which code for proteins that require higher eukaryotic post-translational modifications e.g, cell surface receptors.
- 5) To express intron-containing genomic sequences which cannot be transcribed correctly into mRNA in bacteria or lower eukaryotes.
- 6) To investigate structure-function relationships by analysing the properties of wild type and mutant proteins.
- 7) To identify DNA sequence elements involved in the control of eukaryotic gene expression.

The choice of which mammalian expression system to utilise depends, in part, upon the following five parameters;

- 1) The species and type of mammalian cell available.
- 2) The size of the gene which is to be expressed.
- 3) Whether the experiment requires the transient or stable expression of the protein.
- 4) The method of transfer of the foreign gene.
- 5) The presence of controlling elements in the DNA.

Not all mammalian cells can be transfected or infected efficiently with foreign DNA and cell-lines from different species will not necessarily process the protein as found in the native cell. For example, Chinese hamster ovary (CHO) cells which are routinely used for bulk protein production, tend to add more terminal sialic acid residues to secretory and transmembrane glycoproteins than do human, monkey or mouse cells (Sambrook *et al.*, 1989). This is an important consideration when the expressed protein is intended for human use.

The size of the gene to be expressed within a particular cell-line is important due to the strict packaging size some mammalian viral vectors possess. Therefore, direct DNA transfection may be the only option.

The isolation of stable cell-lines which express the desired protein can be a laborious task, but is generally necessary if more than a few micrograms of protein is required. However, for a number of experiments such as the determination of protein localisation by immunofluorescence, transient expression of the protein is probably all that is required.

The presence of controlling elements for the gene/cDNA is critical to successful expression in that cDNAs can only be expressed if they are placed within a vector which possesses expression elements such as; promoter, enhancer, splice acceptor and/or donor sequences, and polyadenylation signals (see figure 10). Sequences from genomic DNA may possess these required sequences, but there is no guarantee that they will work in the chosen cell-lines available, especially if the DNA sequences are from genes which are expressed in a tissue specific manner, and that the protein is produced at an appropriate level.

2.1 / Current vectors for gene transfer into mammalian cells.

Generally, vectors for introduction of foreign DNA into mammalian cells can be broadly categorised into 2 types: those based on virus infection and those which are based on direct DNA transfer (although sometimes the two systems merge together, as detailed in the text to follow).

2.1.1 / Viral Vector Systems.

Many viruses which infect mammalian cells have evolved mechanisms to specifically increase the production of viral proteins by "hijacking" the cellular transcription/translation machinery. With the increase in the understanding of the molecular biology of many mammalian viruses, has come the ability to genetically engineer viral genomes by the introduction of desired coding regions for foreign genes. Such introduced genes, under the control of viral expression elements, has made it possible to produce infectious recombinant virus particles which, upon infection, produce high levels of foreign gene expression.

Furthermore, viral-mediated gene transfer provides a convenient and efficient means to introduce foreign DNA into the majority of the recipient cells, whereby nearly 100% of the recipient cells may express the foreign gene product. In addition, for many viruses, viral replication can yield multiple copies of the template DNA/RNA which can then serve to amplify the transcription of the foreign gene. Furthermore, because some viruses possess a wide host-range, viral-mediated gene transfer may allow the convenient introduction of foreign genes into a variety of different cell-types from different species. Individual examples of selected viral vector systems are described below.

2.1.1.1 / Papovavirus expression system.

Papovaviruses (derived from Papilloma, polyoma and vacuolating viruses) are non-enveloped viruses which possess a small DS circular DNA genome of approximately 5Kbp. Of all the members of the Papovaviruses, Simian virus 40 (SV40) and polyoma viruses are the best studied. SV40 virus DNA replicates to a high copy number (>100,000 copies per cell, and has thus been more widely employed for expression studies than the lower copy number polyoma virus) in primate cells whilst polyoma virus replicates in murine cells.

The SV40 viral genome (figure 6) can be divided into two regions, the early region which codes for the transforming antigens (large T and small t) and the late region which codes for the structural proteins (VP1, VP2 and VP3). After infection by SV40, the early genes are expressed first, with DNA replication, which is primarily mediated through the Large T antigen, being initiated 12 hours post-infection. Thirty six to forty eight hours post-infection the capsid proteins are expressed, with virion release and cell-death occurring 72-96 hrs post-infection.

Recombinant SV40 viruses can be constructed by the insertion of foreign DNA into the early or late regions of the viral genome (figure 8). Recombinant viral stocks can be produced by providing either a helper virus or a helper cell-line which complements the missing function ,e.g. Cos cells (Gluzman, 1981). Much success has occurred through the use of late replacement recombinants where the cDNA is incorporated in place of the genes encoding the; VP1, VP2 and VP3 proteins.

Figure 8: Schematic depicting the genome organisation of SV40 and its use as a heterologous expression system.

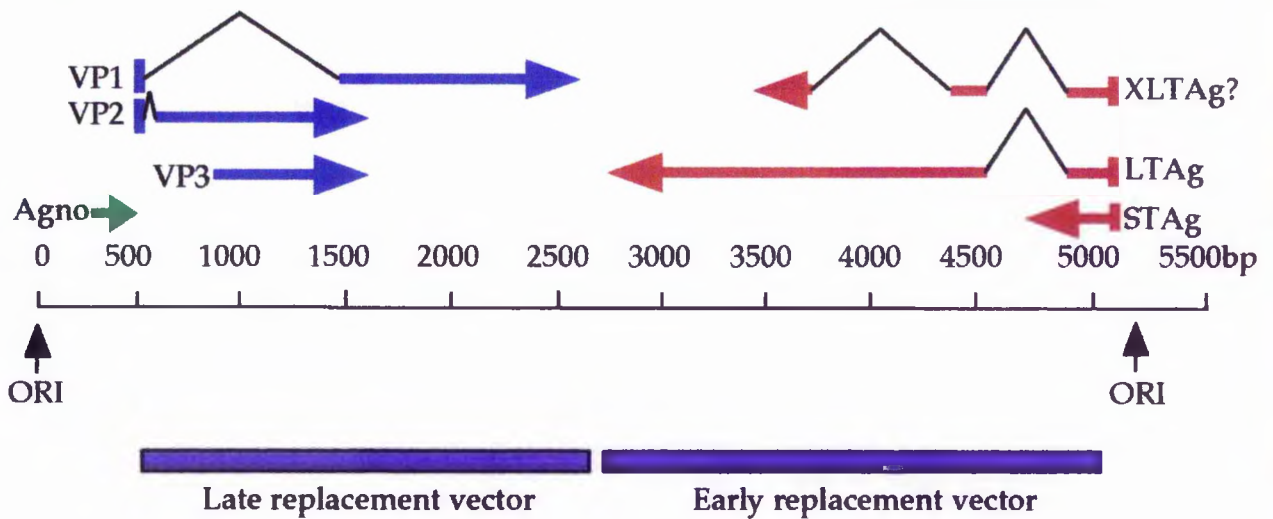
Panel A) *A schematic of the SV40 genome illustrated in a linear form.*

The early gene products, small & large T antigens (and a possible third T antigen: X) are shown in red, and the late gene products (virion proteins (VP) 1, 2 & 3) are illustrated in blue. Also shown are the regions substituted for foreign DNA in the construction of early or late replacement vectors.

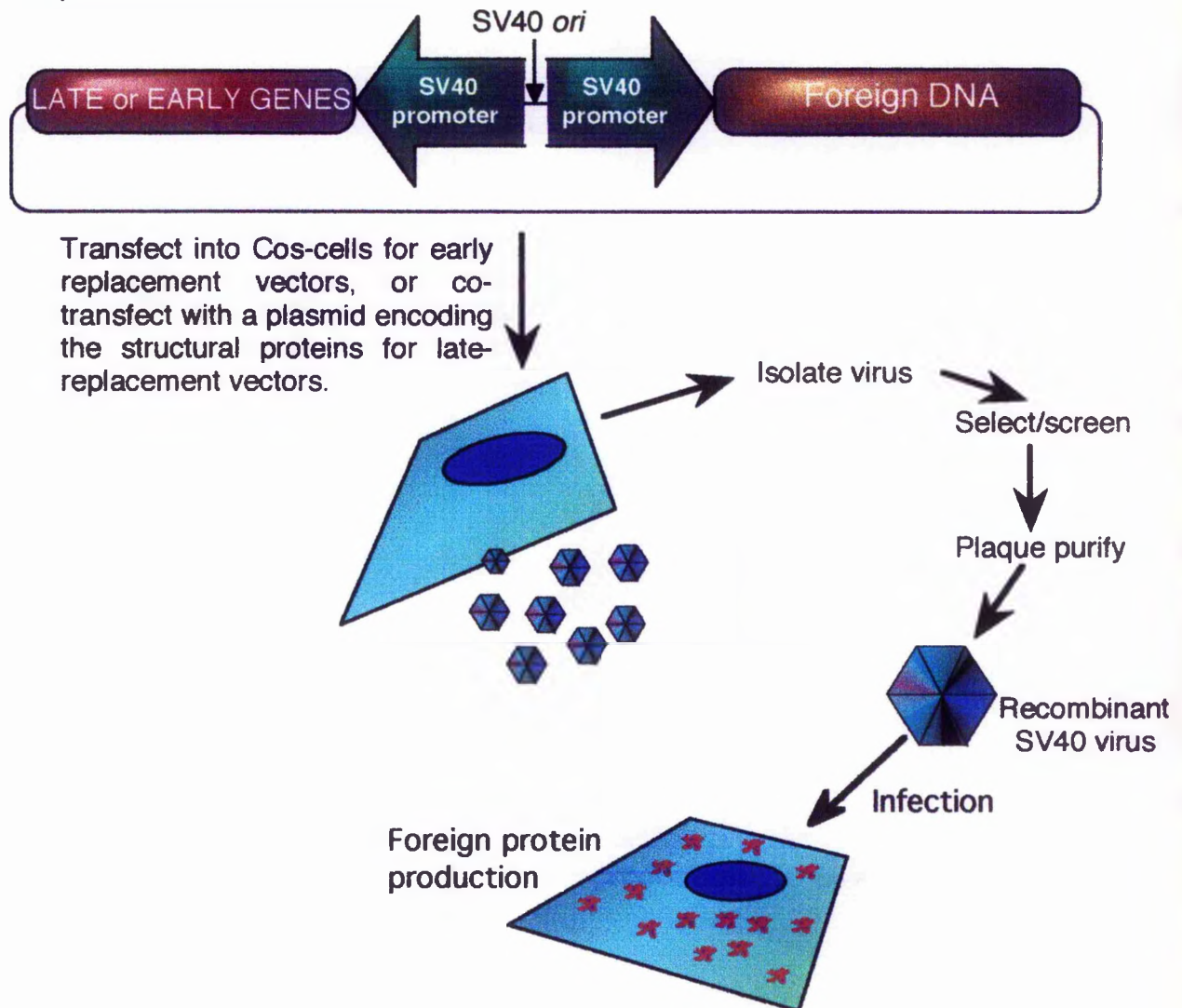
Panel B) *Construction of recombinant SV40.*

Appropriate plasmids are constructed in which the early or late genes of SV40 are replaced by the DNA of choice. Early replacement vectors are transfected into large T antigen expressing cell-lines, such as Cos cells, whereas late replacement vectors are co-transfected with DNA which encodes the virion proteins, or by helper virus. Approximately 72 hrs post-transfection, virus is isolated and screened for foreign protein expression and large titre stocks prepared.

A)



B)



The advantages of the SV40 virus system are the high levels of expression which can be obtained: 1-10 $\mu\text{g}/10^6$ cells, and that potentially toxic proteins can be expressed. The small size of the viral genome also contributes to ease of inserting foreign DNA sequences. The high level of expression observed with the SV40 system is primarily due to high plasmid copy number and the near 100% infection rate. However, there are many disadvantages which have restricted their use, including;

- 1) Primarily only monkey cells can be infected, although infection can occur in human cells, albeit poorly.
- 2) Recombinant SV40 viruses can only be used for transient studies or batch harvest procedures due to cell death 3-4 days post-infection.
- 3) Packaging size restricts the size of the insert to <2.5Kbp.
- 4) DNA rearrangements frequently occur during the replication of these recombinant viruses.

2.1.1.2 / *Vaccinia virus expression system..*

Vaccinia virus is a member of the orthopoxvirus genus of the family Poxviridae, and replicates in the cytoplasm of avian or mammalian cells. Vaccinia virus possesses a large (185Kbp) linear ds DNA genome and encodes its own transcription and RNA processing system. The virus encodes approximately 200 genes: 100 early genes which are expressed soon after infection, and 100 late genes which are expressed after DNA replication at 6 hrs post-infection. Virions are formed about 6 hrs after infection and continue for approximately 48 hrs.

After the eradication of smallpox, vaccinia virus research changed direction and the possibility of inserting exogenous DNA into the virus was investigated, primarily as a means of delivering foreign antigens for immunisation purposes (Panicali & Paoleiti, 1982 and Mackett *et al.*, 1982), although it was observed that vaccinia virus would also be a useful candidate for heterologous protein production.

Recombinant vaccinia viruses are constructed by homologous recombination of plasmid DNA with the viral genomic DNA. This recombination event is usually targeted to viral thymidine kinase gene (other possible sites of integration are listed in Miner & Hruby, 1990), where recombinants are enriched for by the TK⁻ phenotype, or by a colour indicator if the β -galactosidase gene was included on the plasmid (Chakrabarti, 1985). Dominant selectable markers such as neomycin phosphotransferase and *E. coli* xanthine-guanine phosphoribosyl transferase-*gpt*, have recently been utilised (reviewed in Miner & Hruby, 1990).

Early studies with recombinant viruses had the foreign gene being transcribed from the early/late or late viral promoters, 7.5K (Mackett *et al.*, 1984) & 11K (Bertholet *et al.*, 1985) respectively. However, a higher expressing system has been developed, initially by Fuerst *et al.*, (1986) where the specific T7^D RNA-dependant RNA polymerase (T7^D-RDRP) is used to direct transcription of foreign genes which are placed downstream of the T7 ϕ 10 promoter, in place of the above viral promoters, and upstream of the T ϕ termination sequence. Expression of the foreign gene can be achieved in two ways (figure 9);

- 1) By infection of the cell-line with a recombinant vaccinia virus (vTF7-3) which expresses the T7-RDRP, followed by a subsequent transfection with the T7 promoter plasmid (figure 7), or
- 2) By constructing a second recombinant virus which contains the foreign gene downstream of the T7 promoter, and co-infecting the two viruses into the cell-line.

With this approach, nearly 100% of the recipient cells can express the foreign gene. Using the latter system of expression, very high levels of specific mRNA for the foreign gene is produced (24-48 hrs post-infection), often up to 30% of the total cell RNA (Fuerst *et al.*, 1987). However, it appears that only 5-10% of this mRNA is properly capped and thus a significant proportion of message is poorly translated, if at all. To try and alleviate this inherent drawback, a leader sequence from a picornavirus - EMCV virus (encephalomyocarditis virus) - which possesses an internal ribosomal entry site has been added immediately upstream of the AUG initiation codon of the foreign gene, and downstream from the T7 promoter. This

Figure 9: Heterologous protein expression using vaccinia virus (adapted from Miner and Hruby 1990).

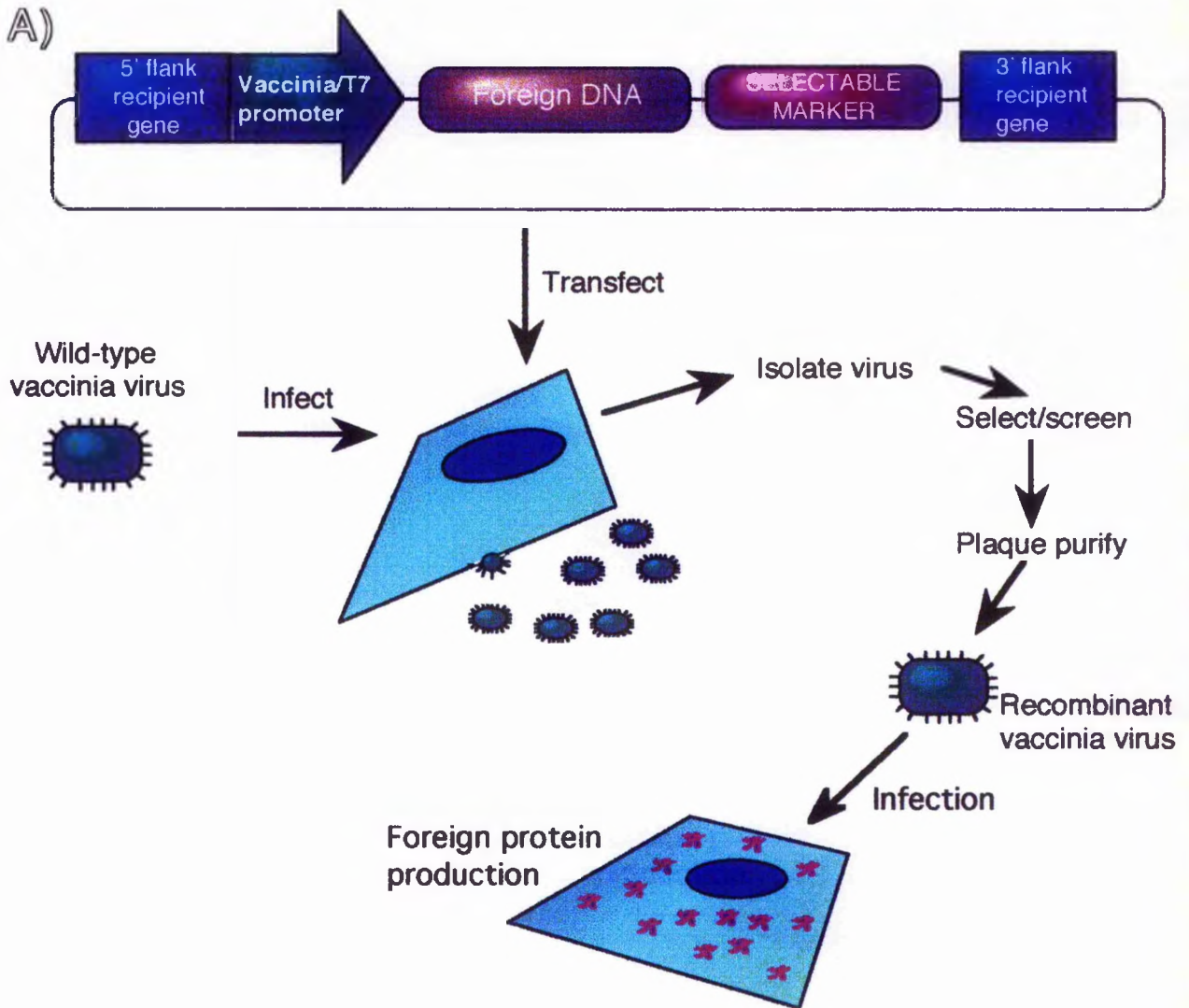
Panel A) *Production of recombinant vaccinia virus vectors.*

DNA of choice is cloned into the appropriate plasmid downstream of a vaccinia virus promoter or T7 promoter. Also present on the plasmid are flanking regions for insertion into the vaccinia genome and elements for the selection of recombinant viruses. The plasmid is transfected into vaccinia virus infected cells and virus is isolated upto 48hrs post-transfection. The isolated virus is selected then screened for recombinant protein production, plaque purified and high titre stocks produced which are used for heterologous protein production.

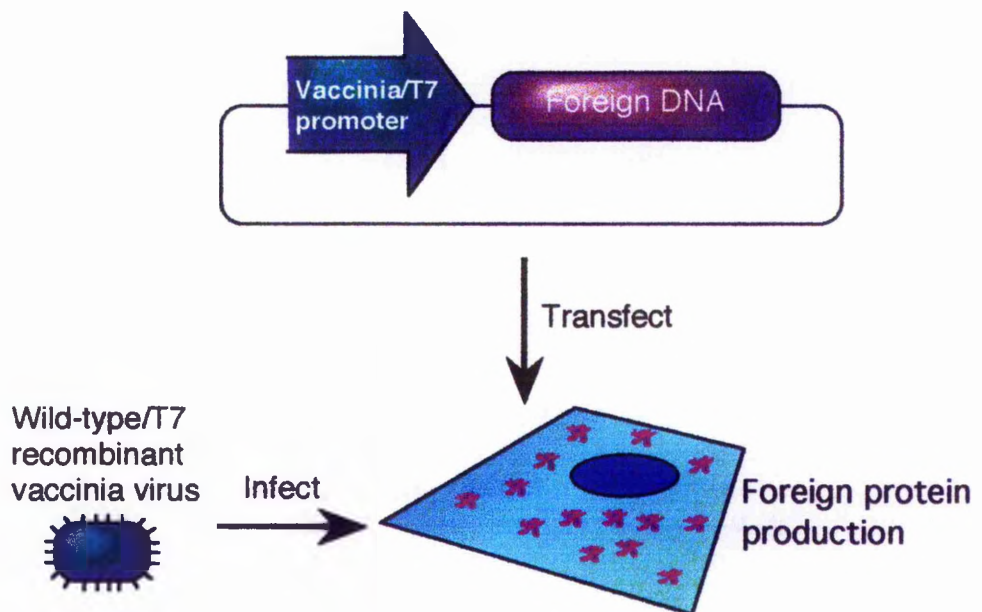
Panel B) *Quick plasmid expression.*

Transient expression can also be achieved by plasmid transfection after infection of the cells with wild type vaccinia virus or a T7 RNA polymerase expressing recombinant vaccinia virus. This alleviates the time-consuming construction of recombinant viruses if only a few experiments are to be performed

A)



B)



addition was shown to increase a reporter gene expression by 4-7 fold, and a further 2 fold if the cells were incubated in hypertonic medium (Elroy-stein *et al.*, 1989).

There are numerous advantages of the vaccinia virus system including;

- 1) Vaccinia is a relative safe virus to work with, especially the attenuated strains.
- 2) Vaccinia has a very wide host range, c.f. SV40 virus.
- 3) Large size of the genome means large genes, or multiple genes can be inserted, up to approximately 25Kbp (Smith & Moss, 1983), c.f. SV40 virus.
- 4) Very high levels of recombinant gene expression can be obtained, even of toxic proteins, especially if the T7 system is utilised.

However, there are of course some disadvantages of using the vaccinia system, including;

- 1) Expression requires an ongoing infection, which ultimately kills the cell, (a highly attenuated avian host-restricted vaccinia virus recombinant has recently been developed by Wyatt *et al.*, (1995) which encodes the T7 RDRP. This avian virus induces a drastically reduced cytopathology when infected into mammalian cells, but still retains the very high levels of recombinant gene expression as observed for the replication-competent vaccinia viruses).
- 2) Vaccinia virus can only be used for transient expressions and batch harvest procedures.
- 3) There is still a small risk of infection and pathological complications, even from the attenuated strains of virus.
- 4) Time consuming to initially produce the recombinant viruses.

2.1.1.3 / Adenovirus expression system

Human adenovirus, of which there are nearly 50 serotypes known, is a member of the mastadenovirus genus of the family Adenoviridae and can cause acute respiratory and ocular disease in humans. The molecular biology of adenovirus has been well investigated (reviewed in Shenk, 1996) and has contributed to the development of adenoviruses as a vehicle for the heterologous expression of recombinant proteins in mammalian cells.

Although adenovirus can be used for the above purpose, most research has been carried out to investigate the use of adenovirus as a delivery system for performing gene therapy and vaccination studies (reviewed in Imler, 1995).

Adenovirus virions are non-enveloped and icosahedral in structure, containing a double-stranded linear DNA genome of approximately 35Kbp which is replicated in the nucleus of primate cells, without the need for host-cell replication/division (Imler, 1995 and figure 10). The genome can be divided into an early (E) and a late (L) region. The early region consists of the; E1, E2, E3 & E4 regions which are essential for transcription & replication of the viral genome, shutting off host-cell protein synthesis and modulating the immune response to the viral infection. The early regions are expressed soon after infection. The late region consists of the; L1, L2, L3, L4 & L5 regions which encode most of the structural proteins of the virion, and are expressed after the onset of DNA replication, 6-8 hours post-infection.

Heterologous protein expression using adenoviruses can be achieved in two ways. By the construction of recombinant viruses (figure 10), or by the virions acting as chaperones to facilitate the uptake of plasmid DNA into cells in a process termed receptor-mediated gene delivery (RMGD).

Adenovirus serotype 5 (Ad5), or its derivatives, is extensively used for the construction of recombinant viruses due to the low pathogenicity and non-oncogenic nature of the virus. Recombinant adenoviruses are constructed, in the majority of cases, by the replacement of viral sequences, primarily the sequences from the early regions: E1 and/or E3. Vectors deleted in E1 are replication-defective and can only replicate in a helper cell-line (293 cells), but expression of foreign genes can still occur in cell culture and *in vivo*. Vectors deleted in the E1 region can accommodate 5.0 - 5.2Kbp of insert DNA. The E3 region is non-essential for virus replication in cell-culture, and deletion of this region can result in the insertion of 4.5 - 4.7Kbp of foreign DNA.

E1-deleted viruses are primarily used for safety reasons and for expression purposes, where DNA inserted into the E1 region has been found to be highly expressed from associated promoters (especially the hCMV IE enhancer/promoter, Xu *et al.*, 1995) and

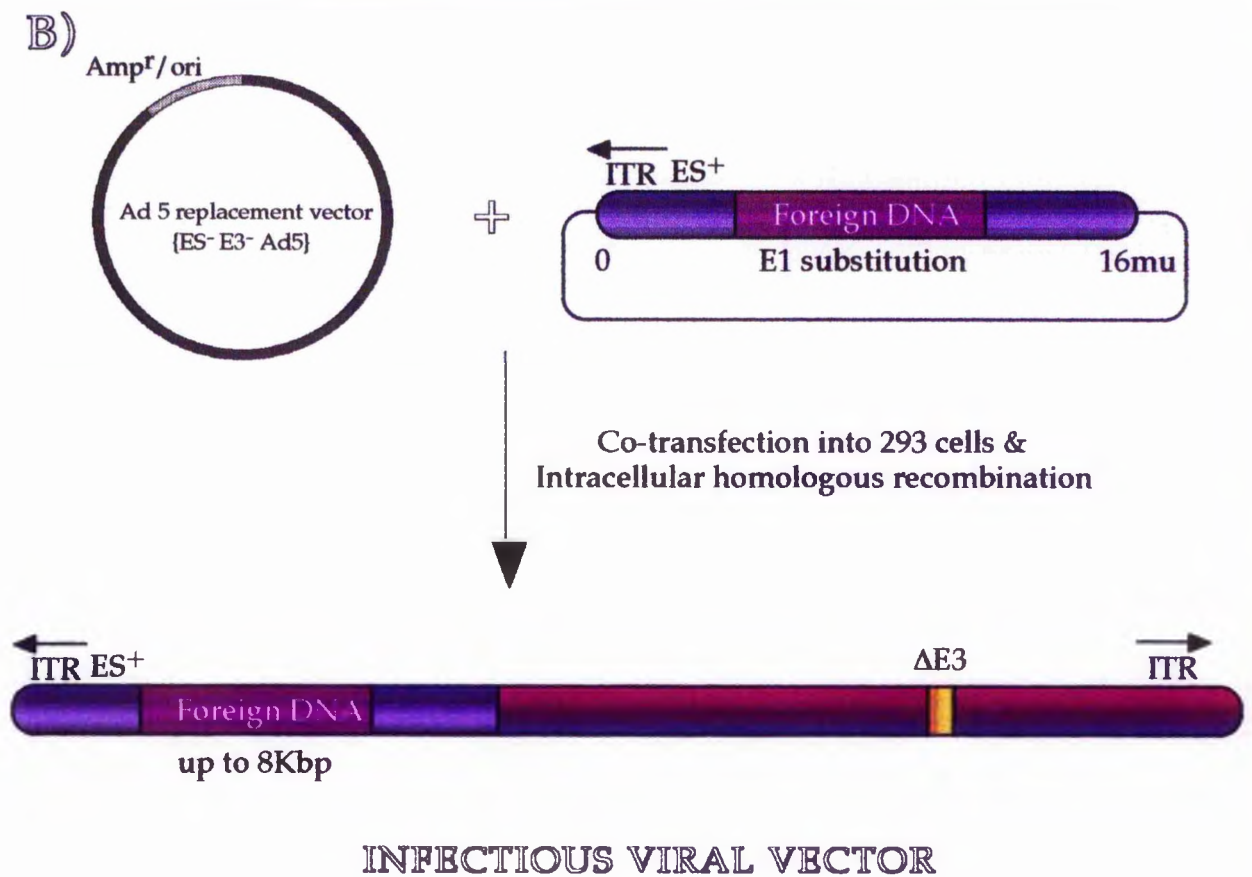
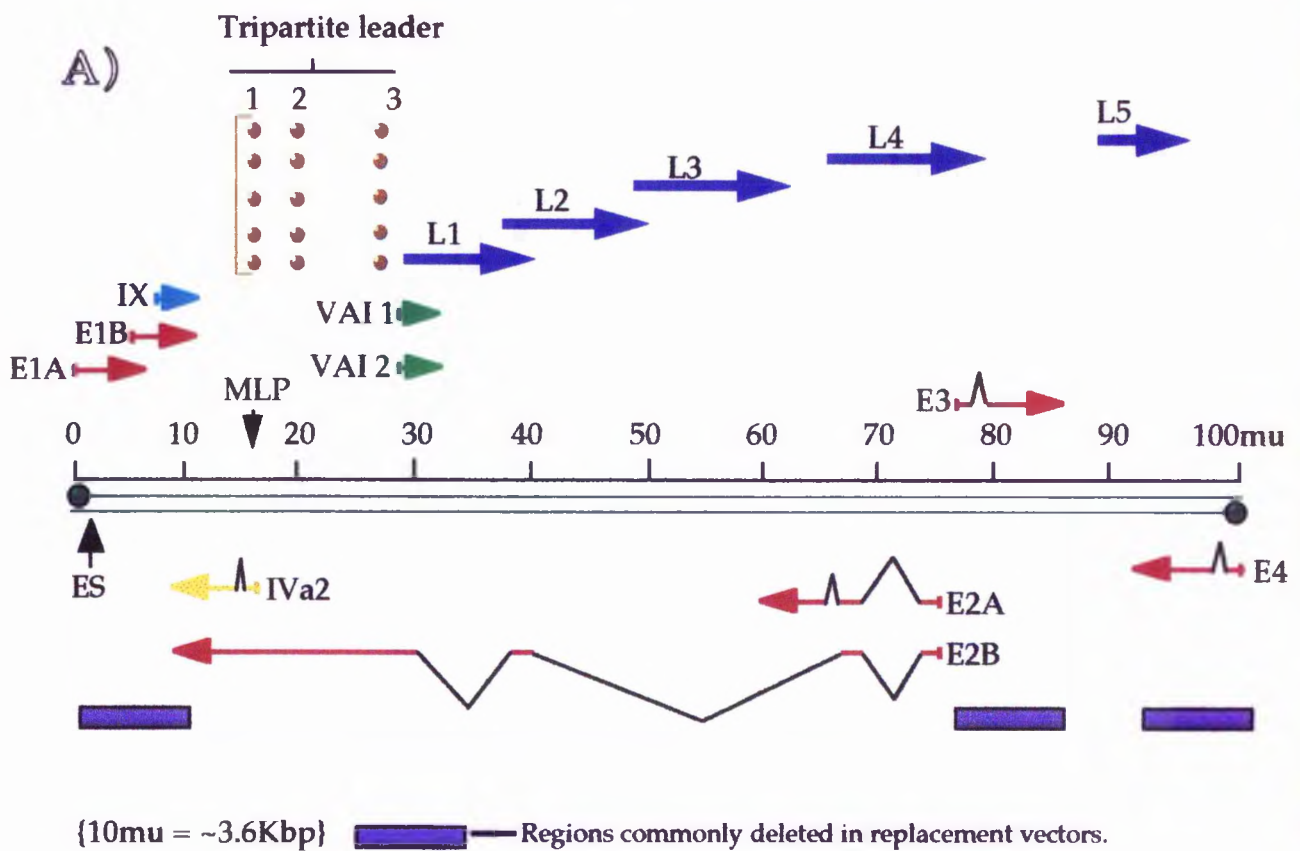
Figure 10: Schematic depicting the genome organisation of adenovirus 5 and its use as a heterologous expression system (adapted from Imler 1996).

Panel A) *A schematic of the Ad 5 genome.*

Illustrated are the major transcription/translation products. Early gene products (E1-E4) are shown in red, late gene products (L1-L5) are shown in blue. Also shown are the positions of the major late promoter (MLP), the encapsidation signal (ES), and regions which are routinely deleted in the construction of recombinant viruses.

Panel B) *A schematic depicting the construction of a recombinant adenovirus expression vector.*

Typically the foreign DNA is inserted into an *E. coli* plasmid shuttle-vector which contains homologous DNA (including an encapsidation signal) to the adenovirus genome flanking the foreign DNA insert. This plasmid is then co-transfected into cells (usually 293 cells, if the E1 region has been deleted) with an ad 5 replacement vector which does not contain an encapsidation signal and is deleted for one of the non-essential ad genes, e.g E3 or E1. Intracellular homologous recombination occurs between the two plasmids, resulting in the production of infectious recombinant virus which is screened for foreign protein production, plaque purified and high titre stocks produced.



independent of the insert orientation (although expression levels have been observed to be up to 7-fold higher if the transcription cassette is in the parallel orientation (Hitt *et al.*, 1995). This is in contrast to inserts within the E3 region, where expression is highly orientation dependant, and associated-promoter independent (reviewed by Graham & Prevec, 1992). The choice of replication-competent or replication-deficient viruses for the construction of recombinant virus depends upon the intended application. For recombinant protein production and vaccination purposes, short-term high-level protein expression, as potentially produced by replication-competent viruses, may be the first choice (Hitt *et al.*, 1995 and Imler, 1995).

Advantages of using adenovirus for recombinant protein production include;

- 1) Adenovirus will infect almost any type & species of mammalian cell.
- 2) The small genome size confers ease of introducing foreign sequences.
- 3) Defective viruses which are essentially non-pathogenic are available.
- 4) Very high titres of recombinant virus can be cultured.
- 5) Adenovirus will infect and replicate in non-dividing cells.
- 6) Up to 8.0Kbp can be inserted into E1 & E3 deleted viral genomes.
- 7) Generally, high levels of recombinant protein can be transiently expressed.

Disadvantages of using adenovirus for recombinant protein production include;

- 1) Currently, only up to 8.0Kbp of DNA can be inserted into E1 & E3 deleted viral genomes, although with the construction of appropriate helper cell-lines, a theoretical limit of approximately 30Kbp could be achieved (recently Caravokyri & Leppard (1995) have isolated a 293 cell-line which constitutively expresses the Ad pIX gene and can thus complement for viruses defective in this region. This subsequently increases the cloning capacity to ~9.2 Kbp).
- 2) There is still a small risk of infection and pathological complications.
- 3) Adenovirus can only be used for transient expressions / batch harvest procedures.

- 4) Currently, even the replication-deficient viruses induce some degree of cytopathology in infected cells, although mutants in the penton base (a structural protein) have been demonstrated to induce a reduced cytopathology (Dr. Paul Freimuth, unpublished results).
- 5) It is time consuming to construct the recombinant viruses (although a recent technical improvement by Chartier *et al.* (1996), whereby homologous recombination between the two plasmids can occur within *E. coli*, will significantly increase the speed of recombinant virus production).
- 6) Recombination between the integrated E1 region in 293 cells and the viral genome can result in the rescue of replication-competent viruses which can contaminate recombinant virus preparations, although new E1 helper cells have been constructed which appear to alleviate this problem (Imler *et al.*, 1995).

As stated earlier, adenovirus virions can be used to enhance DNA delivery into target cells. The natural route of Ad infection in cells requires the disruption of the endosomal membrane, in a pH-dependant manner, with the subsequent release of the Ad virions into the cytoplasm. Uptake of naked plasmid DNA into transfected cells results in a large proportion of the molecules not reaching the nucleus due to degradation within endosomes, therefore the endosomolytic activity of Ad viruses has been utilised to enhance entry of plasmid DNA into target cells.

Plasmid DNA can be attached to a ligand for a specific cellular receptor via a polylysine region present on the ligand. Mixing of high levels of Ad particles (3000 to 10,000 virions per cell) with the plasmid-conjugate mixture has been shown to increase DNA expression by 350-fold with greater than 90% transient transfection efficiency, compared to plasmid-conjugate alone. This result is dependant on cell-lines and Ad strains employed. Moreover, if the plasmid-conjugate is physically linked to the Ad particles via a biotin-streptavidin linkage, expression can be further increased 3-20 fold, again dependant upon the criteria mentioned above (Cotten *et al.*, 1993).

2.1.1.4 / *Adeno-associated virus (AAV) expression system.*

AAV is a parvovirus and is a member of the genus *Dependovirus* belonging to the family Parvoviridae. Parvoviruses possess a linear ssDNA genome of approximately 5Kbp in length. AAV, being a member of the *Dependovirus* genera is, in the majority of situations, replication defective and is thus dependant upon co-infection of the cell either by adenovirus or herpesvirus with subsequent replication of the AAV genome in the nucleus. Under conditions where helper viruses are not present, AAV can integrate, by non-homologous recombination via terminal repeats (tr) present on the AAV genome, specifically into human chromosome 19 where it remains silent until rescued by a helper virus. AAV can establish a stable latent state with a high frequency: 19-70% of exposed cells, even in non-dividing cells (reviewed in Berns, 1996). AAV carries two sets of functional genes and encodes 7 polypeptides: 3 structural proteins (VP1-3) and 4 non-structural proteins (rep proteins) involved in replication of the genome and transactivation of the structural gene promoter.

AAV can be used as a vector for heterologous protein expression by production of recombinant virus particles, or by direct DNA transfection of plasmid DNA. Recombinant viruses are constructed by the replacement of the DNA encoding the structural proteins with the foreign DNA, thus retaining the *rep* genes and supplying the structural proteins, in trans, by transfection of an appropriate plasmid. This procedure allows approximately 2Kbp of foreign DNA to be inserted with the foreign DNA still being integrated at a specific site on chromosome 19. Another approach is to add AAV tr onto both ends of the foreign DNA allowing up to approximately 4.5Kbp of foreign DNA to be packaged into recombinant particles, when both the rep and capsid proteins are supplied in trans (Lebkowski *et al.*, 1988). Expression of the introduced gene has been successfully driven by the AAV promoter 40 (p40, Tratschin *et al.*, 1985) or by other viral promoters such as the CMV IE promoter and the SV40 early promoter.

Advantages of AAV as a vector for heterologous protein expression include;

- a) The small genome size allows ease of handling.
- b) Isolation of stably expressing cell-lines with a known site of integration can be produced (although integration sites outwith chromosome 19 frequently do occur).

- c) AAV can infect most mammalian cells.
- d) AAV has not been associated with any known human disease.
- e) No cytopathology is observed in cells infected with AAV.

Disadvantages of AAV as a vector for heterologous protein expression include;

- a) Difficult to obtain high titre stocks of virus.
- b) Production of virus stocks generally requires the presence of helper viruses.
- c) Only approximately 4.5Kbp of foreign DNA can be packaged (although if the liposome-based transfection method is utilised, larger inserts will be possible).

2.1.1.5 / *Alphavirus expression system.*

Alphaviruses are a group of Togaviruses which infect many cell-types and species, ranging from mosquito to mammalian & avian cells. The genome consists of ssRNA of message sense which is approximately 12 Kbp in length, and encodes two sets of functional genes: the replicase & transcriptase genes (4 in total) and the structural genes (4 in total). All the replication and transcription functions are carried out by the viral replicase/transcriptase in the cytoplasm of the infected cell (Berglund *et al.*, 1996 and figure 11).

Semliki Forest virus (SFV) and Sindbis virus (SV) are two of the alphaviruses which have been intensively investigated for their use as heterologous expression vectors, both for recombinant protein production and vaccination purposes (Berglund *et al.*, 1996). SFV vectors have been utilised to express a variety of different proteins including; human transferrin receptor, chick lysozyme, *E. coli* β -galactosidase and the mouse dihydrofolate reductase (Liljesstrom & Garoff, 1991).

Alphavirus vectors are generally constructed in two forms, replicative recombinants or suicide recombinants. Replicative recombinants have the foreign DNA inserted in addition to the viral genes such that the RNA is infectious and virions are produced. However, these constructs have been shown to be frequently unstable during passage and, in most cases, the production of new virions is undesirable. Therefore, a preferred, and more commonly used approach, is to replace the structural genes with the foreign DNA.

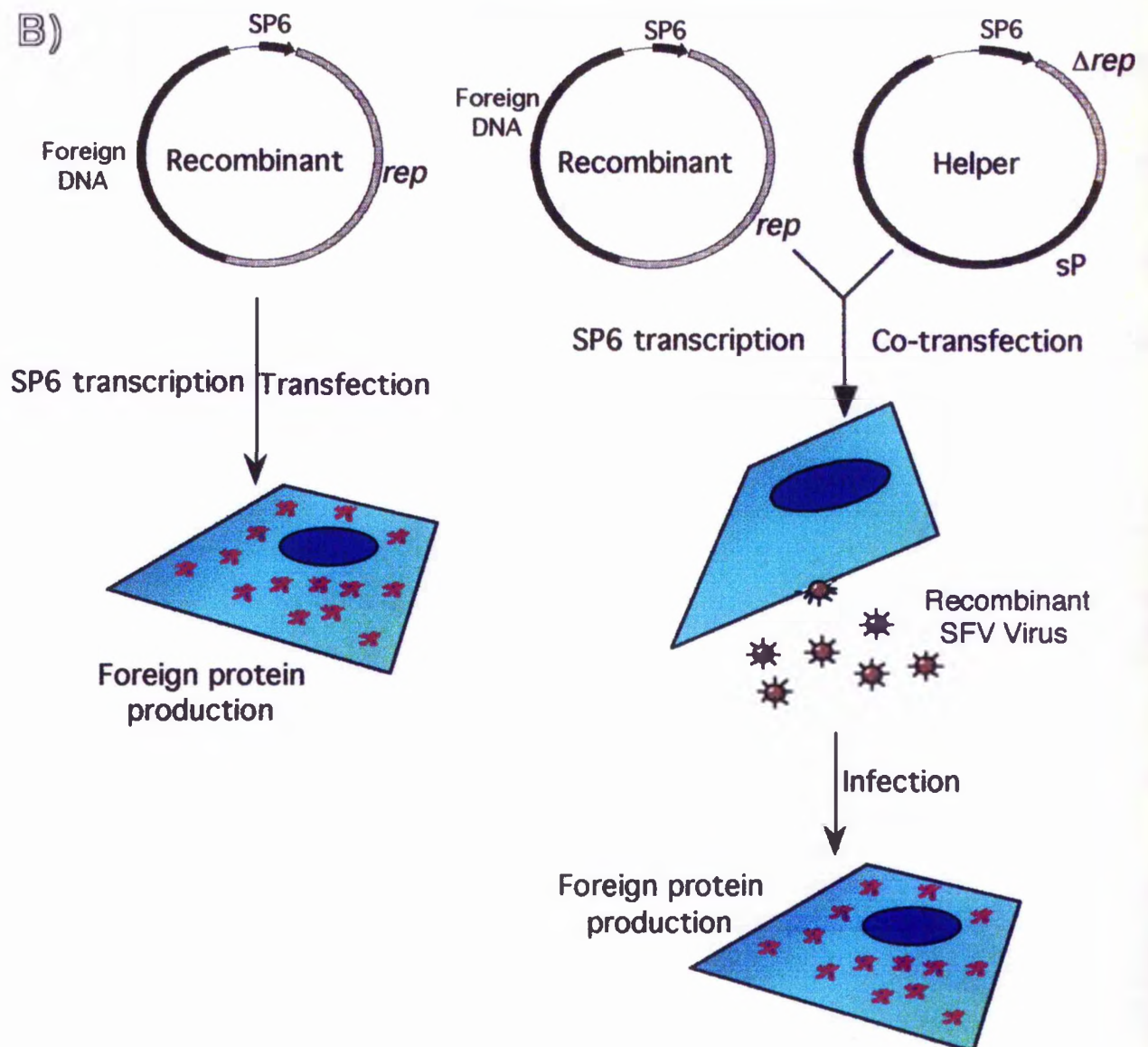
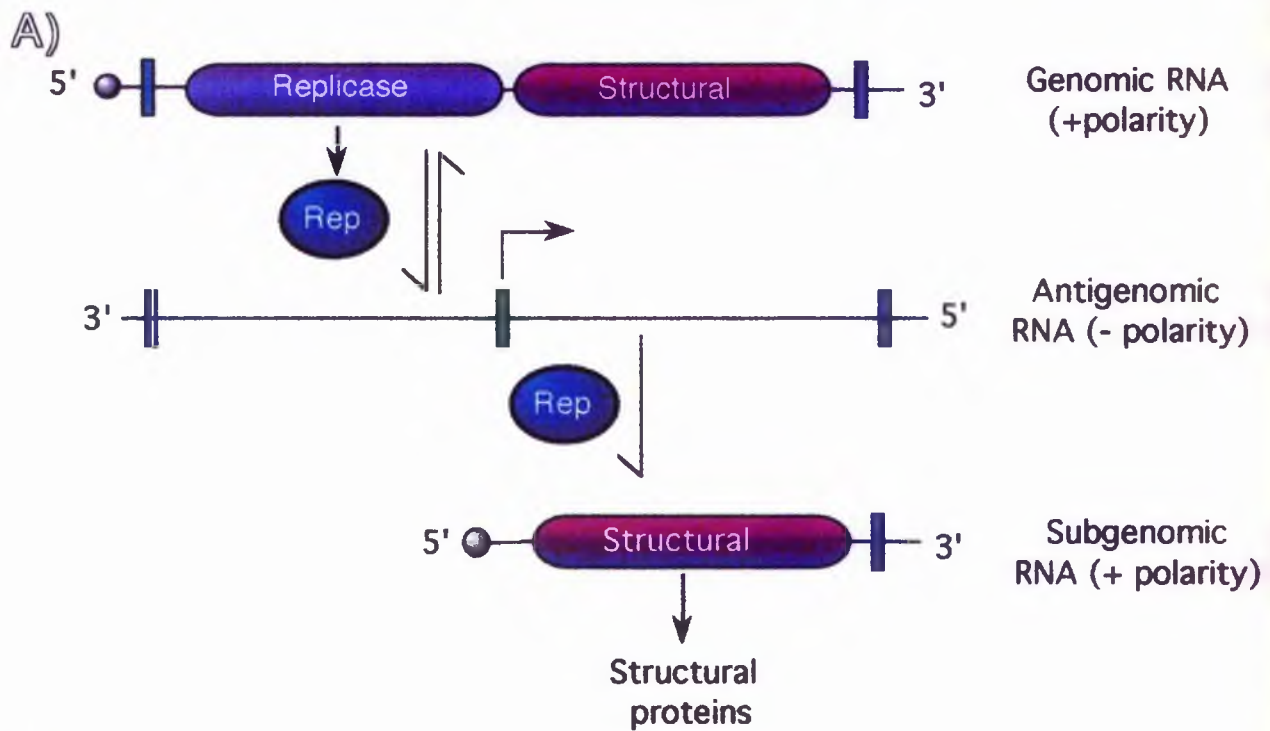
Figure 11: Schematic depicting the replication/transcription of alphaviruses, and their use in the heterologous expression of proteins (adapted from Berglund 1996).

Panel A) Replication/transcription of a typical alphavirus.

Upon entry of the capped (grey circles) message sense ssRNA genome into the cytosol, translation of the *rep* genes produces the replication/transcription (Rep) proteins. The 5' and 3' ends of the genome and anti-genome contain the replication signals (sequences) required for Rep binding (black rectangles). The Rep proteins produce further genomes via a negative polarity ssRNA intermediate, and also transcribes the structural genes, via an internal promoter, which encode the virion proteins.

Panel B) Heterologous protein expression using an alphavirus system.

Typically, foreign DNA is inserted into a plasmid which contains only the *rep* genes from an alphavirus downstream from an SP6 promoter. Following linearisation and *in vitro* transcription, the message sense RNA is transfected into appropriate cells and subsequent amplification of the RNA, via the Rep proteins, results in high level translation of the foreign gene. Alternatively, a helper plasmid which has deleted *rep* genes, but which encodes the structural proteins, can be co-transcribed/transfected resulting in the production of recombinant alphavirus particles which can be used for infection of cells resulting in the production of the protein of interest.



DNA of interest is cloned into appropriate plasmid-based vectors which then, generally, serve as templates for the *in vitro* synthesis of recombinant mRNA. The mRNA is then transfected into the cell-line of choice and subsequent translation of the replicase/transcriptase genes leads to the amplification of the transfected RNA (up to 50% of total cellular mRNA) with subsequent large amounts of protein being produced for up to 80 hrs post-transfection (~1mg of LacZ per 10^7 cells, i.e up to 25% of the total cellular protein), (Liljestrom & Garoff, 1991). Due to inherent differences between cell-types in the uptake of RNA, a packaging system has been developed whereby recombinant RNA is packaged into infectious virions via co-transfection with packaging-deficient helper RNA molecules which encode the structural proteins (figure 9).

A third approach for the expression from alphavirus-based vectors has recently been developed by Dubensky *et al.* (1996) involving a layered DNA-RNA vector system. This system, which is independent of helper vectors, involves the use of a eukaryotic promoter (e.g. CMV I.E. promoter) which drives the expression of the replicase/foreign gene after transfection of the plasmid DNA into cells. The replicase then takes-over the replication/transcription of the RNA in the cytoplasm.

Recently, it has been demonstrated that the addition of DNA, encoding the first 30 amino acids of the capsid gene, to the 5' end of the foreign DNA sequence results in an increase in expression of the fusion protein of up to 10-fold (Liljestrom, 1994). Furthermore, incubation of the infected/transfected cells in medium containing elevated levels of K^+ ions has been observed to result in a further increase in protein expression of up to 5-fold (Liljestrom, 1994).

In summary, the main advantages of the alphavirus expression vector system include;

- a) Very high levels of protein production can be achieved.
- b) Replicase/transcriptase complex caps the mRNA leading to efficient translation.
- c) RNA is in the positive sense, i.e infectious.
- d) High titres of recombinant viruses can be obtained without the need for time consuming recombination, screening and plaque purifying techniques.
- e) Vectors can be RNA, DNA or viral particle based.

- f) Alphaviruses have a wide host-range.
- g) Very little morphological change occurs for up to 48hrs post-infection/transfection.
- h) SFV and SV have not frequently been associated with human disease.
- i) For biosafety purposes, helper viruses and possibilities of recombination to produce WT virus can be avoided by using the DNA-based plasmid approach. Furthermore, attenuated strains are available which require the exogenous addition of protease to cleave the viral spike protein to induce infectivity.

Disadvantages of the alphavirus expression vector system include;

- a) Currently can only be used for transient and batch harvest procedures.
- b) Safety concerns for working with alphaviruses and the possibility of recombination events occurring between helper RNA and vector RNA (see above).

Two other viral vectors which are being increasingly investigated for use in gene therapy studies are members of the retroviruses and herpesviruses. These viruses are currently rarely used for heterologous protein production, but can be useful as a means of infecting cells which are impervious to infection by other viruses or by the various methods of naked DNA transfection. Reviews on herpesvirus vectors are given in Glorioso *et al.*, 1995; Fink *et al.*, 1996 and Leib & Olivo, 1993, and for retroviral vectors in Levinson, 1990.

2.1.2 / Direct DNA Transfer.

Viral-mediated-gene-transfer (VMGT) for heterologous protein production has, in a large number of situations, been extremely useful. However, due to the aforementioned problems associated with VMGT, direct DNA transfer is still widely used as a means for transient protein production and as the initial stage to the production of stably expressing cell-lines.

In designing a plasmid-based vector for the expression of a foreign gene there are numerous factors which have to be considered, including;

- a) Choice of promoter element to initiate transcription.
- b) Choice of enhancer element to enhance transcription.

- c) Whether an intron is required for efficient transcription/RNA processing.
- d) Choice of transcription termination and polyadenylation sequences.
- e) Is an optimal translation initiation signal present?
- f) Is the codon sequence optimal for mammalian cells?
- g) Is the protein to be transiently produced, or are expressing-cell-lines necessary?
- h) If cell-lines are to be produced, is the gene to be integrated into the cellular genome or would an episomal maintenance be beneficial?
- i) If integration is required, is random integration sufficient or would site-directed integration be advantageous?
- j) If cell-lines are produced, is expression of the gene to be constitutive or regulated?
- k) Choice of selectable marker if cell-lines are produced: amplifiable vs non-amplifiable?

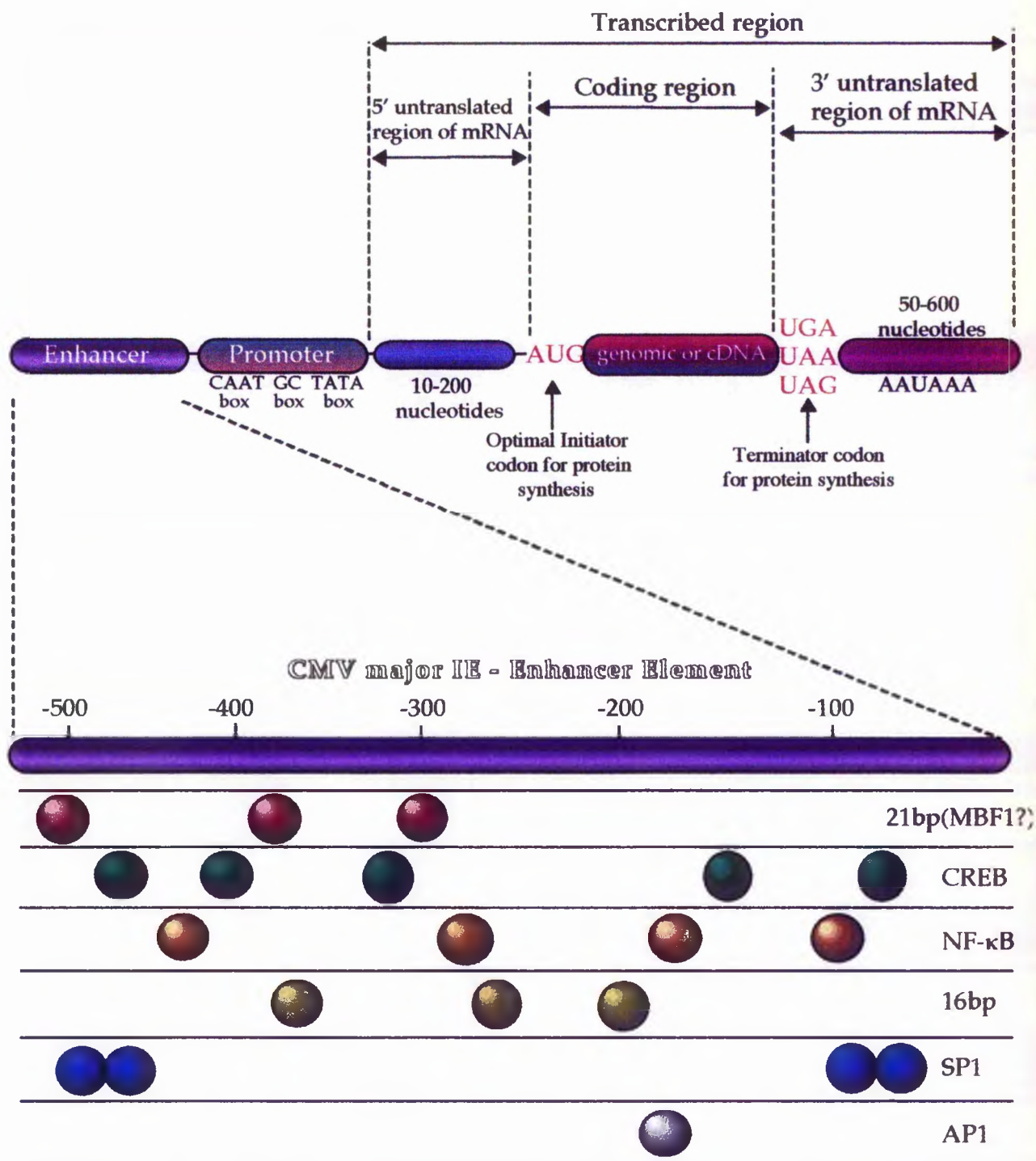
A detailed account of promoter/enhancer elements and the factors which interact with them will not be given here but can be found in books such as *Eukaryotic Gene Transcription* (Goodburn, 1996).

2.1.2.1 / Choice of promoter and enhancer.

A eukaryotic promoter is a cis-acting element of approximately 100bp in length which is present immediately upstream of a gene and possesses sequences required to assimilate RNA polymerase II (RNA pol II) and various transcription initiation factors for the initiation of gene transcription (figure 12). Enhancers are also cis-acting elements which, when bound by trans-acting factors can serve either to enhance or repress transcription from associated genes (figure 12). The key properties which make a DNA sequence an enhancer element include; (1) they tend to be relatively large elements and may contain repeated sequences which can function independently, (2) they can influence gene expression over considerable distances from the gene, (3) they may function in either orientation, (4) they may function in a position-independent manner being upstream or downstream from the gene, (5) they may function in a cell-type or tissue-specific manner, and (6) they appear to increase the probability but not the level of gene transcription (Walters *et al.*, 1995).

Figure 12: Diagrammatic representation of an RNA Polymerase II transcription unit.

One of many possible arrangements of sequence elements in genes coding for eukaryotic mRNA molecules is illustrated. Shown is the enhancer region from the human cytomegalovirus (hCMV) major immediate-early promoter and some of the transcription factors which bind to consensus sequences within the enhancer (see text for more detail).



Promoters for driving heterologous gene expression can be divided into 2 categories: constitutive and inducible. The majority of promoter/enhancer elements used for the heterologous constitutive expression of proteins in mammalian cells are those derived from viruses. Examples include the; CMV major immediate early enhancer/promoter (CMV IE), SV40 early/late promoters, RSV long terminal repeat (RSV LTR), adenovirus major late promoter (Ad MLP) and the HSV thymidine kinase promoter (HSV TK). These elements are often used due to their activity in a wide variety of cell-types and the high level of gene expression they confer.

Generally it has been observed that the human CMV IE element confers the highest level of expression of genes transfected into primate cells followed by the RSV LTR then the SV40 promoters (Wilkinson & Akrigg, 1991; Kronman *et al.*, 1992; Harrison *et al.*, 1995 and Rotondaro *et al.*, 1996). However, of course this is not absolute and is largely dependant upon the chosen cell-line for expression of the gene. For example, Rotondaro *et al.*, (1996) found that the murine CMV major IE element performed 2-fold better in murine cells compared with the human version of the promoter. The high level of gene expression observed with the CMV IE element is possibly reflecting the ubiquitous nature of the transcription factors which bind to the element such as; NF- κ B, CREB and AP1, which is present in many cell types (Wilkinson & Akrigg, 1991 and figure 12). Moreover, expression from the CMV IE element can be further increased by employing cells which express the adenovirus E1a gene product (which transactivates transcription from the CMV enhancer), such as 293 and CHO L761h cells.

Furthermore, various mammalian enhancer/promoter elements have also been utilised with some limited success, such as the; immunoglobulin enhancer/promoter, β -actin promoter/enhancer, γ -globulin and β -globulin enhancer/promoters, elongation factor 1 α promoter, and human muscle-specific promoters (Dahler *et al.*, 1994; Kawamoto *et al.*, 1988; Mizushima & Nagata, 1990 and Harrison *et al.*, 1995). However, some of these elements appeared to be very tissue-specific which can be a major advantage or drawback depending upon the intended application.

With the aforementioned success of the vaccinia virus/T7 RDRP system, attempts have also been made to develop a non-viral T7 expression system which would thus not be

dependant upon host-cell availability of transcription factors and RNA polymerase. Cell-lines permanently expressing the T7 RDRP (located either in the nucleus or the cytoplasm) have been developed (reviewed in Lieber *et al.*, 1993) which apparently expressed CAT and human growth hormone to levels similar to that achieved with the hCMV major IE promoter. However, this high expression was due to illegitimate activation of the T7 promoter by RNA pol II and although attempts to produce mutant promoters which were exclusive for the T7 RDRP have been partially successful, expression levels have been no greater than those obtained with the RSV LTR or hCMV major IE promoter (reviewed in Lieber *et al.*, 1993).

Inducible expression systems have been widely utilised in bacteria and yeast and have recently been developed for use in mammalian expression studies. Inducible promoters can be divided into two groups: (1) promoters which are responsive to endogenous cellular transactivators, and (2) promoters which are regulated by exogenous bacterial regulatory proteins. Examples of the former group include the; interferon β -promoter, heat-shock promoter, metallothionein promoter and the mouse mammary tumour virus promoter, with induction from these promoters being in the range of 5-200 fold (reviewed in Levinson, 1990). However, these promoters tend to produce high levels of basal expression and their induction may cause pleiotropic effects to the cells. Furthermore, as with the example of hormonal control, induction may require the presence of specific receptors on or within the cells.

Therefore, recent interest in the use of exogenous prokaryotic-based control elements has increased dramatically. Such mediation of expression can be achieved in two manners: the transrepression principle and the transactivation principle. Two inducible prokaryotic repressor/operator (R/O) elements which have been utilised to regulate gene expression in mammalian cells, via the repression principle, are the *E.coli* lactose (*lac*) operon (where isopropyl- β -D-thiogalactopyranoside, IPTG, is the inducer) and the (Tn10 - derived) tetracycline (*tet*) operon (where tetracycline is the inducer) (Gossen *et al.*, 1994). The basis of the regulatory effect is the close proximity of the respective operator sequences to the eukaryotic promoter, such that binding of the repressor protein to the operator sequences inhibits the formation of an active RNA pol II transcription complex at the promoter.

However, such Lac R/O-based systems have never proved satisfactory with induction been slow, incomplete and usually requires near cytotoxic concentrations of IPTG. The corresponding Tet R/O system has been developed in yeast and plant cells, but for reasons not fully understood, but perhaps due to low intracellular concentrations of the tet repressor, has not been adapted to mammalian cells (Gossen *et al.*, 1994).

Therefore, the second approach - the transactivation principle - has been developed whereby transcriptional activation domains are fused (via recombinant DNA technology) to the repressor proteins of the above systems. Great success has recently been achieved using the tetracycline transactivator (tTa) protein which is a fusion protein between Tet R and the acidic domain of herpes simplex virus VP16 (Gossen & Bujard, 1992 and figure 13). This tTa protein stimulates transcription from a promoter sequence, usually the minimal hCMV major IE promoter, when combined with upstream *tet* operators. Transcription is repressed in the presence of tetracycline and activated in the absence of tetracycline and can be regulated over a 10⁵-fold range (Gossen & Bujard, 1992). A recent development in the tTa system has been the isolation of a reverse TetR phenotype where the presence of tetracycline derivatives, especially doxycycline, are required for binding of the reverse tTa (rtTa) molecule to the operator sequences and thus for activation of transcription (Gossen *et al.*, 1995 and figure 13). This system should have great implications for regulation of gene expression in transgenics and of control of expression of cytotoxic genes. The advantage of the tTa/rtTa system over similar Lac-based systems is the high affinity of tetracycline for TetR/rTetR (1000-fold greater than IPTG is for LacR) which results in lower intracellular concentrations required for regulated gene expression (Gossen & Bujard, 1992). Furthermore, many tetracycline derivatives are available and their pharmacokinetics have been studied in great detail which will facilitate the application of these systems to transgenic animals, or appropriate cell-lines *in vitro*.

2.1.2.2 / Intron present or absent, and choice of 3' untranslated region (UTR).

The presence of an intron has been found to be either essential or detrimental to the expression of a gene depending upon factors such as; (1) promoter/enhancer element used to

Figure 13: Diagrammatic representation of tetracycline-controlled transactivator (tTa) systems (adapted from Didcock, 1996).

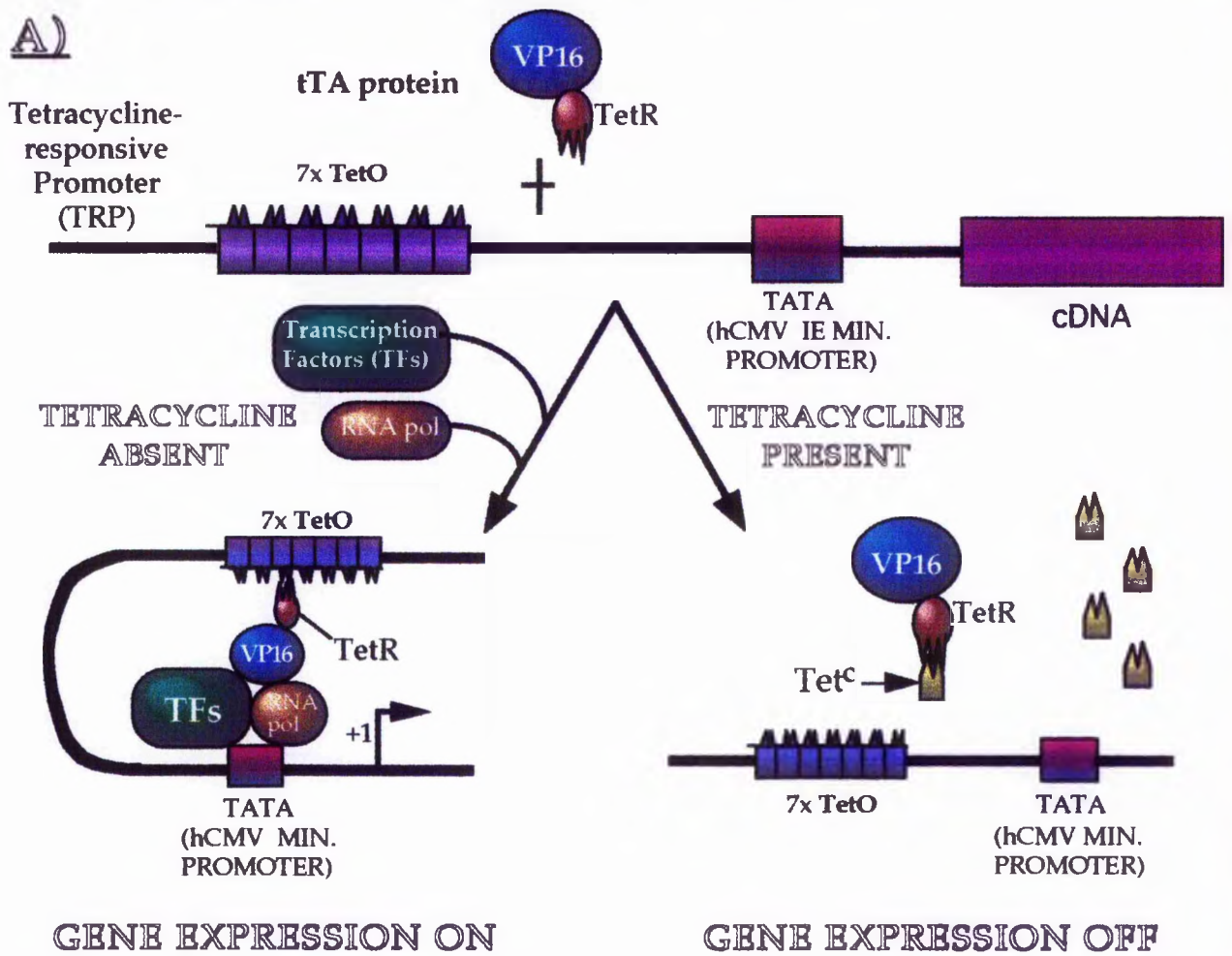
Panel A) *Transcriptional repression in the presence of tetracyclines.*

In the presence of tetracycline, the TetR domain of the tTa protein does not bind to the upstream tet operator sequences and thus transcription of the downstream DNA is minimal. In the absence of tetracycline, tTa can bind to the upstream tet operator sequences and, through the action of the acidic domain from the herpes simplex virus VP16 protein, recruit various transcription factors to upregulate the expression of downstream genes.

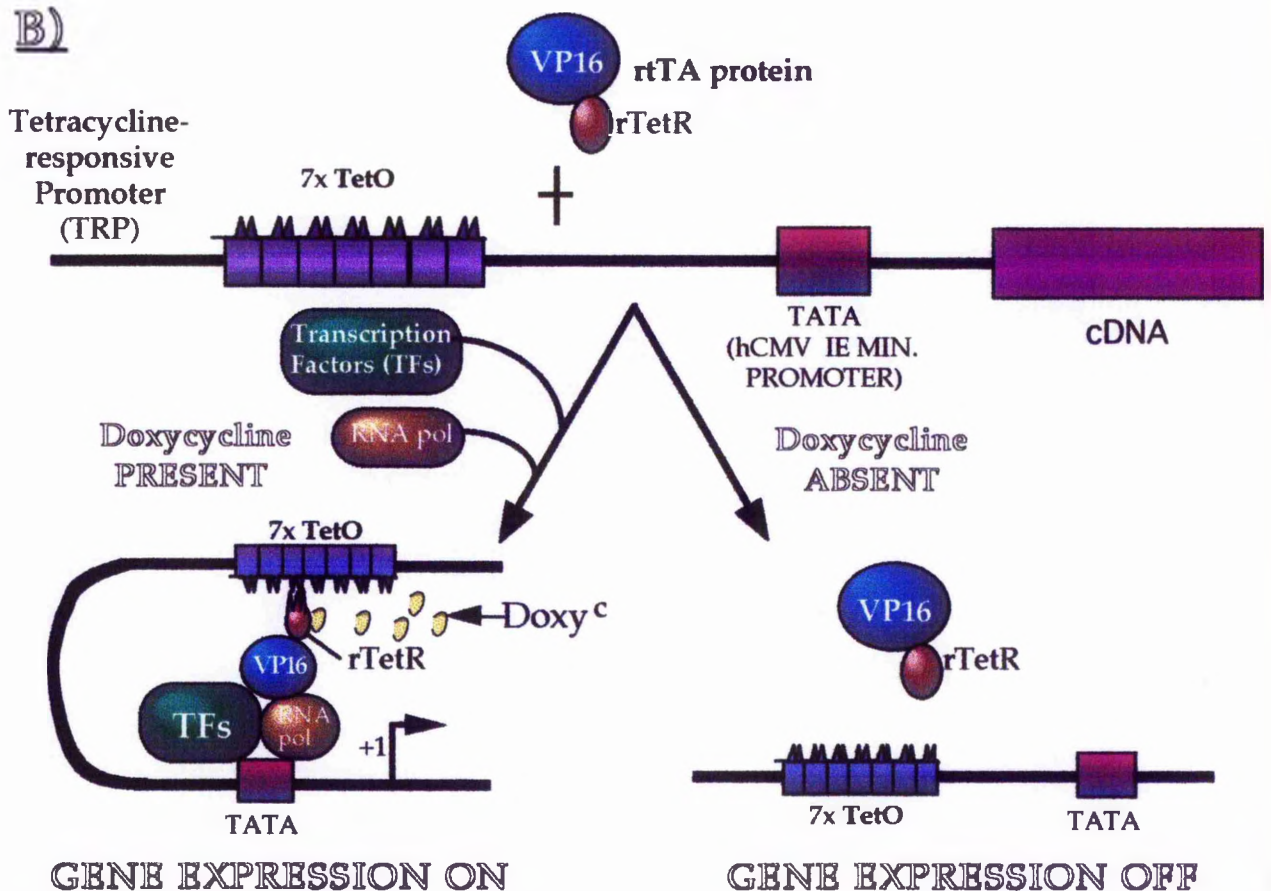
Panel B) *Transcriptional activation in the presence of tetracyclines.*

A mutated form of the tTa protein-rtTa- binds to the tet operator sequences in the presence of tetracyclines (especially doxycycline) and thus transcription of downstream genes is activated. In the absence of doxycycline, rtTa cannot bind to the tet operator sequences and therefore transcription of downstream genes is minimal.

A)



B)

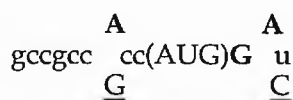


drive transcription. (2) the gene which is being expressed, and (3) cell-line in which the gene is being expressed. It has been observed that for genes driven by the immunoglobulin enhancer/promoter an intron is required for efficient expression although the requirement was not restricted to a particular intron (Kriegler, 1990). An intron is not apparently required for efficient expression from the CMV IE element (Nueberger & Williams, 1988) although some workers have found evidence contradictory to this. Furthermore, it has been shown by Buchman & Berg (1988) that expression of the β -globin gene from an SV40 early promoter required the presence of an intron in which case transcription was increased by over 400-fold. Moreover, experiments performed by Petitclerc *et al.* (1995) demonstrated that the SV40 late promoter was more efficient than the SV40 early intron, although the greatest gene expression was observed when a chimeric intron, composed of an adenovirus donor & immunoglobulin acceptor sites, was employed. Other genes such as the dihydrofolate reductase gene, which do not possess introns, showed a modest increase in expression of 10-20 fold if an intron was included in the construct. This intron-dependant gene expression was also noted for the HSV *tk* promoter. In contrast to expression in mammalian cells *in vitro*, there is a profound requirement for the presence of introns for expression of introduced genes in transgenic animals (Brinster *et al.*, 1988).

The 3' untranslated region (figure 12) contains the signals for transcription termination and subsequent addition of poly A residues. The choice of polyA signal can have a significant effect on the expression of the chosen gene due to the role in mRNA stability provided by the polyA signal. For example ornithine decarboxylase mRNA has a half-life of only $1/2$ hour compared to dihydrofolate reductase mRNA with a half-life of nearly 100 hours (reviewed in Wolfe, 1993). Efficient signals for polyadenylation frequently used for heterologous expression are derived from the; SV40 late polyA signal, bovine growth hormone poly A signal, mouse/human β -globin polyA and the CMV IE gene polyA due to the long half-life that they confer to the mRNAs produced. Furthermore, Wilkinson & Akrigg (1991) have demonstrated that using a polyA signal combined with the respective promoter, in this case the CMV IE promoter and polyA signal, can increase expression by 7-fold than when the same promoter was used in conjunction with the SV40 late polyA signal.

2.1.2.3 / *Optimal translation initiation sequence.*

The efficiency with which a scanning mammalian ribosome recognises an AUG codon has been shown to be highly dependent upon the context of the surrounding sequences and is not simply the first AUG that is encountered. From studies carried out by Kozak (1989) and Grunert & Jackson (1994) the optimal translation initiation codon is:



where residues in bold capital letters exert the most positive influence. Those residues in underlined capital letters exert a modest positive influence, with the remaining lower case residues exerting a yet smaller but nevertheless positive influence. For heterologous protein expression in mammalian cells, initiator codons should conform to this optimal context if high levels of protein production is required.

Furthermore, some genes can contain upstream AUG codons in the 5' untranslated region which should ideally be removed when constructing the vector. Similarly, the presence of significant secondary structure near the initiator AUG can significantly affect translation and thus most vectors should be designed to contain 5' untranslated sequences derived from the particular promoter element used for the transcription initiation, or by incorporating leader sequences such as those derived from the β -globin gene or adenovirus tripartite leader.

2.1.2.4 / *Optimisation for mammalian-cell codon usage.*

tRNAs specifying a particular amino acid can recognise multiple codons present in the coding sequence of DNA due to the degeneracy of the genetic code. The abundance of tRNAs which recognise particular codons vary considerably between organisms, and this can lead to a significant reduction in protein production if particular tRNAs are present in low abundance. Altering the codon usage for optimal recognition by mammalian tRNAs has been shown to significantly increase the amount of HIV gp120 and green-fluorescent-protein by up to 45-fold (Haas *et al.* 1996 ; Levy *et al.*, 1996 and Zolotukin *et al.*, 1996).

2.1.2.5 / *Transient or stable expression of the protein.*

For many purposes, transient high-level expression of a particular protein is all that is required, and one of the most common vector systems for such expression is that derived from plasmids which contain the SV40 virus *ori*, and which are subsequently transfected into SV40 LTA_g expressing cells such as Cos-cells. The LTA_g amplifies the plasmid to >10,000 copies per cell resulting in high-level protein production in the order of 1-10 μ g per 10⁶ cells (Levinson, 1990). The LTA_g can also be encoded on the same plasmid eliminating the need for using Cos-cells.

This system, and other transient systems, may be further enhanced for protein expression, for certain plasmids, by the addition of 2-aminopurine to the medium or by co-transfecting a plasmid encoding the adenovirus VAI genes. Both of these treatments enhance expression by inactivating the dsRNA-activated inhibitor kinase (DAI kinase, also known as PKR, p68, p1 or dsI) which is activated by the presence of dsRNA, and results in a repression of translation (Kaufman & Murtha, 1987; Akusjarvi *et al.*, 1987; Tiwari *et al.*, 1988 and Tanigawa *et al.*, 1993).

However, for many purposes, the establishment of expressing cell-lines is required and suitable vectors must be utilised for this purpose. Establishment of expressing cell-lines requires either the transcription unit to be integrated into the cellular genome or to be maintained in an episomal form in the cell nucleus. Integration into the cellular genome is a rare event and thus plasmids must contain a suitable marker gene, under control of an appropriate promoter element, for both selection and maintenance of the expressing cell-lines.

Such marker genes can be amplifiable or non-amplifiable. Non-amplifiable markers include the commonly used neomycin aminoglycoside phosphotransferase (*neo^r*), hygromycin B phosphotransferase (*hyg^r*) and histidinol dehydrogenase (*hisD*). Commonly used amplifiable markers in which the gene copy number within the genome can be significantly increased (to over 1000-fold) include dihydrofolate reductase, P-glycoprotein 170 (multidrug resistance gene) and glutamine synthetase (reviewed in Levinson, 1990 and Kane & Gottesman, 1993).

However, amplification of a gene can be time-consuming, involving several weeks work, and does not always lead to a significant increase in protein production even if the gene copy number has been significantly increased. Other methods for obtaining high-level expressing clones include, the tedious screening of hundreds of unamplified stable clones by classical techniques, and by the more "user-friendly" fluorescent activated cell sorting (FACS) if the gene of interest (or a marker gene) is cell-surface expressed and antibody is available.

Due to integration occurring rarely, and since positional effects can occur when DNA integrates in a random manner, much interest has been directed towards establishing site-directed integration vectors. Such vectors have been developed either for homologous recombination (widely used in embryonic stem cell technology for the generation of transgenic mice) or site-specific recombination, in which the latter relies upon exogenously added recombinase (principally from bacteriophage and unicellular yeasts) derived from two major families: the integrase family and the resolvase/invertase family (reviewed in Sauer, 1994). Frequently used recombinases are the simple Cre recombinase from bacteriophage P1 and the FLP recombinase from *S. cerevisiae*, both of which do not require additional host-specific co-factors for activity.

However, at best, recombinase-mediated integration appears only to be as efficient as illegitimate recombination, (albeit, when recombination did occur it was, in the majority of cells, at the chosen site (Fukushige & Sauer, 1992)) and also necessitates the introduction of appropriate target sites in the genome to begin with, usually via homologous recombination. This poor performance is, in part, due to poor introduction of the recombinase into the cells. However, Anton & Graham (1995) have constructed an adenovirus vector which expresses the Cre recombinase and have demonstrated promising initial results using this strategy.

An alternative approach to using integrating vectors for the construction of expressing cell-lines is by utilising vectors which possess sequences/genes required for the episomal maintenance of the vector in the host-cell nucleus. Advantages of using such vectors, over integrating vectors, is that; (1) the frequency of producing expressing cell-lines is increased, (2) positional effects from integration are reduced, (3) generally the gene will have a moderately high copy number, (4) the episome can be easily recovered from the cell-

line (useful for expression cloning; Levinson, 1990; Margolskee *et al.*, 1988 and Shen *et al.*, 1995) and (5), maintenance as an episome facilitates the undertaking of replication and mutagenesis studies.

Sequences/genes for the maintenance of an episomal state are usually derived from viruses whose lifecycle includes the autonomous replication of its DNA in an episomal form. Examples of viruses whose sequences/genes have been utilised for producing episomal vectors are; SV40, BK virus (a human papovavirus), Bovine papilloma virus (BPV), and Epstein-Barr virus (EBV). Some of the properties of these vectors are depicted in table 3.

Vector-System	Host cell	Stable Copy No.	Oncogenic Proteins Produced	Vector Size *1	Strict Rep ^{ln} once per Division	Low mutation frequency
SV40 Virus	Primate	~750 per cell	Yes	~ 5Kbp	No	No
BK Virus	Primate	1-300 per cell	Yes	~11Kbp	No	No
BPV Virus	Rodent Canine Porcine	1-250 per cell	Yes	~ 9Kbp	No	No
EBV Virus	Primate Canine Rodent *2	1-50 per cell	No	~11Kbp	Yes	Yes

Table 3 - Episomal-based vector systems for the production of expressing cell-lines.

*1 - Approximate size of a eukaryotic/prokaryotic shuttle vector before insertion of the gene of interest.

*2 - Can replicate in rodent cells if human sequences are incorporated into the vector. This however increases the vector size to ~29Kbp (Krysan & Carlos, 1993).

SV40 has had little use in the production of expressing cell-lines due to the high copy number of the vector in the nucleus which frequently leads to cell death, and due to the required presence of the LTA_g, an oncogene, which could limit studies of other potential oncogenes. However, Gerard & Gluzman (1985) have successfully isolated an SV40-based cell-line, in which the vector is maintained as an episome (~750 copies per cell), which inducibly expresses the influenza virus haemagglutinin protein.

As with SV40, BK virus-based vectors code for an oncogenic protein, which has limited their use. However, the episome copy number is lower (up to ~250 copies per cell (Gianni *et al.*, 1988)) than with SV40 and thus facilitates the isolation of expressing cell-lines. Human interferon- β 1 and HSV glycoproteins G-1 & G-2 have been successfully expressed in human cell-lines using BK virus-based vectors (Gianni *et al.*, 1988 and Sabbioni *et al.*, 1995).

More frequently used are BPV-derived vectors. These vectors are based upon the complete BPV genome or a 5.5Kbp transforming fragment and efficiently transform rodent cells (but do not produce viral particles), and are maintained at up to ~300 copies per cell (Sambrook *et al.*, 1989). However, due to the transforming proteins produced, the high frequency of DNA rearrangement which can occur, and the lack of replication in primate cells, their use has not been extensively widespread.

Episomal vectors based on EBV have become popular and many commercial EBV-based vectors are available. Episomal replication of EBV vectors requires the presence of a *cis*-acting region from EBV termed *ori* P (~1.7 Kbp in length, Yates *et al.*, 1984 and Sugden *et al.*, 1985), and a *trans*-acting virally encoded factor EBNA-1 (Epstein-Barr nuclear antigen 1) which binds specifically to the *ori* P region and non-specifically to genomic DNA (Lupton & Levine, 1985 and Yates *et al.*, 1985). Numerous cDNAs have been successfully expressed from EBV-based vectors including human interferon- γ (Young *et al.*, 1988), influenza virus HA (Jalanko *et al.*, 1989) and adenovirus pIX (Caravokyri & Leppard, 1995). A further advantage of EBV-derived vectors is the possibility of their use in gene therapy/DNA vaccination studies. They are episomally maintained, replicate once per cell cycle, possess a low mutation frequency, and do not encode an oncogenic protein vital to their episomal maintenance. Indeed, recently the EBNA-1 protein has been shown to be inefficiently presented to MHC 1 restricted cytotoxic T lymphocytes due to the presence of the gly-ala repeats within EBNA-1 which generate a *cis*-acting inhibitory signal (Levitskaya *et al.*, 1995).

CHAPTER 4

THE MAMMALIAN-CELL EXOCYTOTIC PATHWAY.

1 / Introduction.

Proteins destined for subcellular organelles, the cell-surface or for secretion into the external milieu, possess specific signals in their primary amino acid sequence which are recognised by the appropriate cellular machinery for the successful targeting of the protein. In the last 20 years, an enormous amount of information has become available, on both the signals and the machinery involved in protein trafficking within the mammalian cell. Apart from the inherent scientific curiosity of elucidating the mechanisms of protein trafficking, a major beneficiary of this knowledge will be the biotechnology industry, where the targeting of proteins within the mammalian cell can be altered by the addition/deletion of the appropriate signals. A recent example of such technology is in the development of intracellular targeting of antibodies to subcellular compartments for therapeutic benefit (Richardson & Marasco, 1995). Furthermore, it has recently become apparent that diseases such as progeria, which result in premature ageing, may be the result of aberrant protein targeting (Rosenblum *et al.*, 1996). A diagrammatic representation of the exocytotic pathway is shown in figure 14.

2 / Targeting of proteins to the ER.

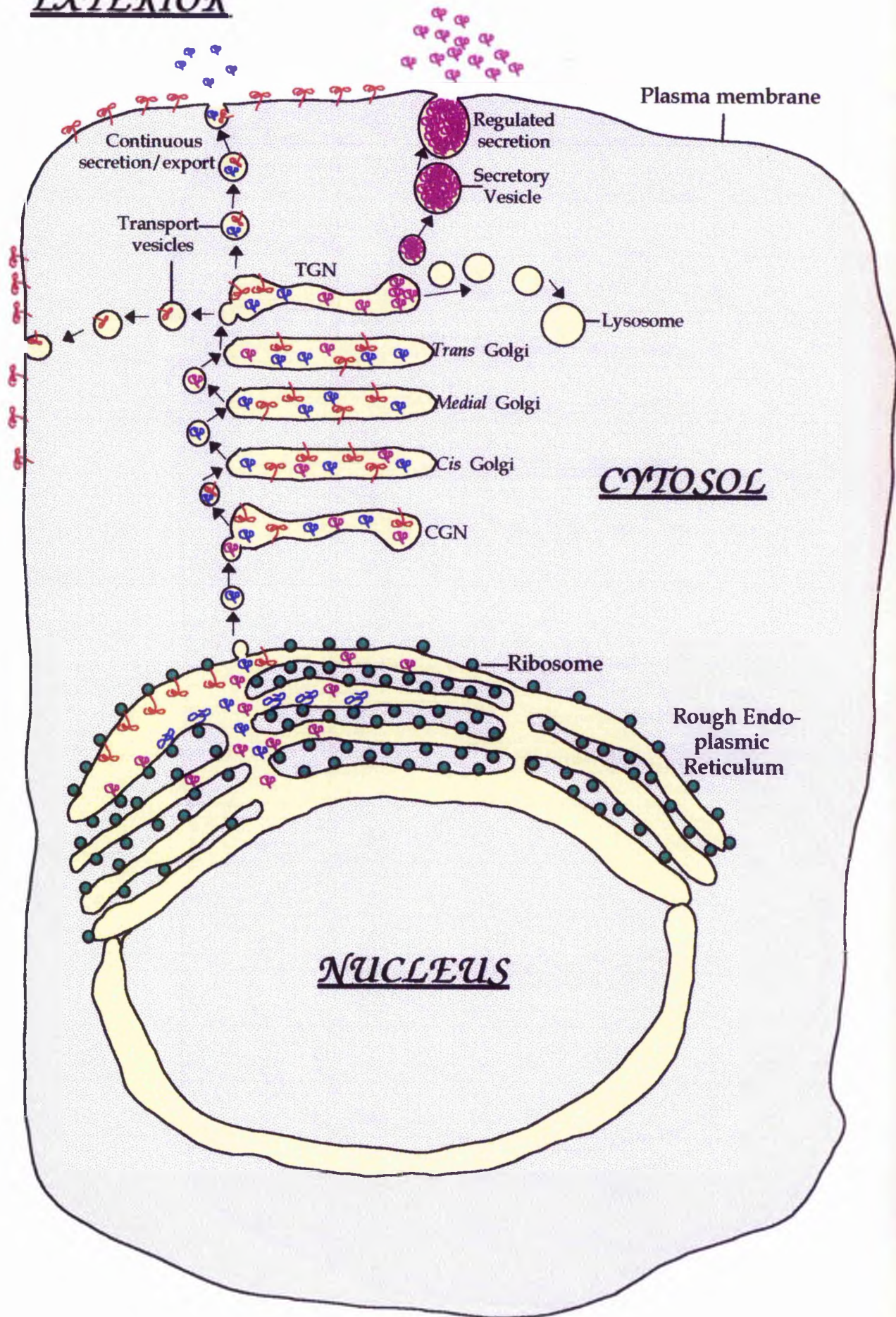
2.1/ Signal sequences.

Proteins for export to the cell surface or external milieu are generally targeted to the ER by the presence of a sequence of about 20-40 amino acids, the signal sequence (SS),

Figure 14: Protein exocytosis from a mammalian cell (adapted from Chidgey *et al* 1993).

Proteins destined for secretion or transport to the plasma membrane are usually initially targetted to the ER. Unless no signals are present within the protein sequence or conformation, then transport of the protein continues through the subcellular compartments of the Golgi complex until the proteins either reach the plasma membrane (if the proteins are membrane anchored) or are secreted into the external milieu. The latter can occur constitutively from cells, or can be, in addition, regulated as found in exocrine and neuronal cells where release occurs via an external stimulus (see text for more detail).

EXTERIOR



present at the N-terminus of the polypeptide. However, there are notable exceptions to this rule where some proteins enter the ER despite lacking a recognisable signal sequence. These include the secreted proteins interleukin 1 α & 1 β and factor XIIIa. The mechanism of this alternative targeting is unknown (but does appear to be cell-type dependent), but may involve ABC transporters similar to the MHC peptide transporters TRAP 1&2 (reviewed in Kuchler, 1993). Furthermore, secretion of certain proteins from mammalian cells, such as basic fibroblast growth factor occurs in an ER-golgi independent process, the mechanism of which is unclear (Mignatti *et al.*, 1992).

The original hypothesis for the presence of an N-terminally located signal for the targeting of nascent polypeptide to the ER was published by Blobel & Sabatini (1971), which was strengthened from experimental evidence compiled by Milstein *et al.* (1972). Lingappa *et al.* (1984) produced the first secreted chimaeric protein when they demonstrated the secretion of alpha-globin (normally located in the cytoplasm) from mammalian cells when the N-terminus was altered by the addition of the N-terminal sequence from bacterial β -lactamase. This experiment also demonstrated the ubiquitous nature of the signal sequence from organisms as diverse as primates and bacteria.

Kohara *et al.* (1992) has defined the ER targeting sequence to be composed of primarily 3 regions; an N-terminal positively charged region of 1-7 residues containing 1-3 positive amino acids, a hydrophobic core comprising 6-12 amino acids, with the optimal being 8, and a 4-6 amino acid region c-terminal to the hydrophobic core. By definition this signal is necessary and sufficient to direct proteins to the ER. Upon arrival at the ER membrane the signal can then either be cleaved from the protein (the majority of secretory and some membrane-bound proteins) or remain attached and function as the transmembrane anchor (signal-anchor membrane proteins). Type I membrane proteins are defined as those where the N-terminus is translocated (i.e extra-cytoplasmically located) whilst for type II proteins the C-terminus is translocated. Membrane proteins with a cleavable signal sequence always possess a type I topology, membrane proteins with a signal-anchor can have either a type I or II orientation (reviewed in High, 1995).

2.2/ SRP-dependent targeting.

In 1981, Walter and Blobel discovered the protein complex, present in mammalian cells, which recognised the signal sequence: the signal recognition particle, or SRP. Equivalents of SRP have been discovered in bacteria and yeast but whether they can function co-translationally has not yet been established. The SRP is a cytosolic ribonucleoprotein, which may be pre-bound to ribosomes, and is comprised of 6 polypeptides and a 7S ssRNA. Recent evidence suggests that the interaction of SRP with the signal sequence is not as specific as first postulated and that it is the presence of another cytosolic factor NAC (nascent-polypeptide-associated complex) which binds ubiquitously to emerging polypeptides, thereby preventing them from being bound by SRP, unless a signal sequence is present (Wiedmann *et al.*, 1994; Luring *et al.*, 1995 a & b; Jungnickel & Rapoport, 1995).

The protein subunits within the SRP appear to play 3 roles; (1) signal sequence recognition, as performed by the 54Kda polypeptide, (2) control of translation, as carried out by the 9 & 14Kda polypeptides, and (3) targeting to the ER membrane, as performed by the 68 & 72Kda polypeptides. A further result of the interaction of SRP with the signal sequence is to retain the signal sequence in an unfolded conformation for membrane insertion. The 54Kda protein has been resolved to 12-15 Angstrom (Czarnota *et al.*, 1994) and shown to possess two domains: an M domain which binds the signal sequence and a G domain which binds GTP (binding of the SRP to the SS has been demonstrated to be independent of GTP hydrolysis, Rapiejko & Gilmore 1994). The 7S ssRNA is postulated to act as a "scaffold" in holding all the polypeptides together.

3 / Translocation of polypeptides across the ER membrane.

The SRP/signal-sequence/ribosome/mRNA complex (translocation complex) is recognised by at least three proteins on the ER membrane: the SRP receptor (SR, or docking protein), the signal sequence receptor (SSR) and the putative ribosome receptor (RR, a 180Kda protein). The SRP receptor, which binds GTP, is comprised of two polypeptides; the α -subunit (35Kda) and the β -subunit (22Kda). Upon binding of the translocation complex to the ER receptors, hydrolysis of GTP occurs resulting in the release of the SRP and the

continuation of translation, followed by translocation of the growing polypeptide chain through an aqueous channel present in the ribosome & ER membrane from which the cytosol has been sealed off by the binding of the ribosome to the ER membrane (Crowley *et al.*, 1993). *In vitro* depletion analysis of SSR or SR demonstrated that SR was essential for targeting and translocation of the nascent chain, whilst depletion of SR did not have any observable effects (Migliaccio *et al.*, 1992).

How the polypeptide is translocated across the ER membrane and the proteins involved in this process are not well understood. In yeast and mammalian cells, several ER resident proteins have been identified which are likely to play a role in translocation of some, but not all proteins. These include the; SRP, RR, Sec61 protein (which is composed of 3 polypeptides; α , β and γ), TRAM (translocating chain associating membrane protein), TRAP α (translocon associated protein α), ribophorins and GRP78-BiP (reviewed in Ng & Walter, 1994 and Simon, 1993). The latter protein, and other luminal proteins, may play a role in the later stages of translocation via the folding of the newly translocated polypeptide. The mechanism of translocation is unclear but current evidence suggests the presence of an aqueous pore, which is formed, predominantly, by multi-subunit integral membrane proteins such as Sec61 (Crowley *et al.*, 1993; Jungnickel & Rapoport, 1995 and Lyko *et al.*, 1995).

4 / Modifications of newly translocated polypeptides within the ER.

4.1 / Polypeptide folding and oligomerisation.

Upon translocation of the polypeptide into the ER lumen, the polypeptide may be modified by numerous processes such as; signal sequence cleavage, N-linked glycosylation, disulphide bond formation, folding and subunit assembly. The ER lumen provides an oxidising environment, contains high levels of calcium and resident ER enzymes which are involved in the above processes (reviewed in Gaut & Hendershot, 1993).

The signal sequence may be cleaved by signal peptidase, a multi-subunit complex present on the luminal side of the ER membrane. Signal peptidase is composed of 5 polypeptides, 2 of which are homologous to Sec 11 from yeast. There is no apparent

homology to bacterial leader peptidase (although signal peptidase has been shown to cleave bacterial signal sequences) or to other proteases.

Translocation exposes domains on the protein (especially hydrophobic regions) which will usually be hidden after folding. There are specific proteins in the ER (termed chaperones), such as GRP78-BiP (a member of the heat shock family of proteins), which reversibly bind to the translocating polypeptides due to an ATPase activity which is stimulated by hydrophobic peptides. This binding appears to aid in the folding of the protein by preventing polypeptide aggregation. ER calcium levels may also have a role in protein folding and assembly, via calcium-dependant chaperones such as; GRP94, ERp72 and p88 (also known as calnexin). Calnexin has been shown to bind to the T-cell receptor, MHC class I molecules, membrane-bound Igs and numerous viral glycoproteins such as the vesicular stomatitis virus (VSV) G protein and the influenza virus HA, and aid in their folding and assembly (reviewed in Bergeron *et al.*, 1994; and Hammond *et al.*, 1994). Indeed Hammond *et al.* (1994) have demonstrated that calnexin binds to glucosylated forms of proteins and releases the protein upon trimming of the glucose residues by the resident ER proteins glucosidases I & II. Properly folded proteins are generally not glucosylated and are thus not bound by calnexin, therefore indicating that calnexin acts as a quality control mechanism for protein folding via the addition/removal of glucose residues. Pro-collagen has been shown to interact with the chaperones hsp74, GRP94 & GRP78-BiP but it is not known whether, in this case and others, chaperones act competitively, sequentially or co-operatively (reviewed in Gething, 1991). However, at least for the VSV G protein, it has been demonstrated by Hammond & Helenius (1994) that Bip bound maximally to early folding intermediates of G protein, whereas calnexin bound after a lag period to more folded molecules, thereby illustrating a sequential order for chaperone binding.

The oxidising environment within the ER allows the formation of intra & intermolecular disulphide bonds which aids the stabilisation of the tertiary structure of the polypeptide and, in many cases, allows the formation of homotypic or heterotypic oligomers. Disulphide bond formation is carried out by the enzyme protein disulphide isomerase (PDI)

which is the beta-subunit of the enzyme prolyl 4-hydroxylase. PDI has two catalytic sites which function equally and independently.

4.2 / Addition of carbohydrates to nascent polypeptide chains.

Glycosylation of newly translocating polypeptides can also occur within the ER. Glycosylation may take the form of asparagine-linked oligosaccharides (within the consensus tripeptide sequence, or sequon, asn-X-ser/thr/(cys), where X is any amino acid except proline), termed N-glycosylation, or serine/threonine-linked oligosaccharides, termed O-glycosylation (O-glycosylation of membrane/secretory proteins in mammalian cells, in contrast to yeast cells, does not occur within the ER but occurs exclusively within the golgi complex, Trimble & Verostek, 1995). Not all potential N-linked sequons are utilised, and those which are may not be glycosylated to the same extent as at other sequons on the same polypeptide. This feature appears to be determined by the primary, secondary and tertiary structures of the polypeptides. Furthermore, the carbohydrates which are added onto a polypeptide is also determined by the enzymes present within a cell, which is thus dependent upon the cell type, species and stage of development and differentiation (reviewed in Warren, 1993).

Oligosaccharides are added to the translocating polypeptide via a lipid carrier intermediate (usually dolichol). Saccharides are added sequentially (via integral membrane-bound glycosyl transferases) to the lipid carrier on the cytosolic face of the ER membrane. The completed oligosaccharide chain is then transferred to the appropriate sequon on the luminal side of the ER membrane (Wolffe, 1993, and figure 15).

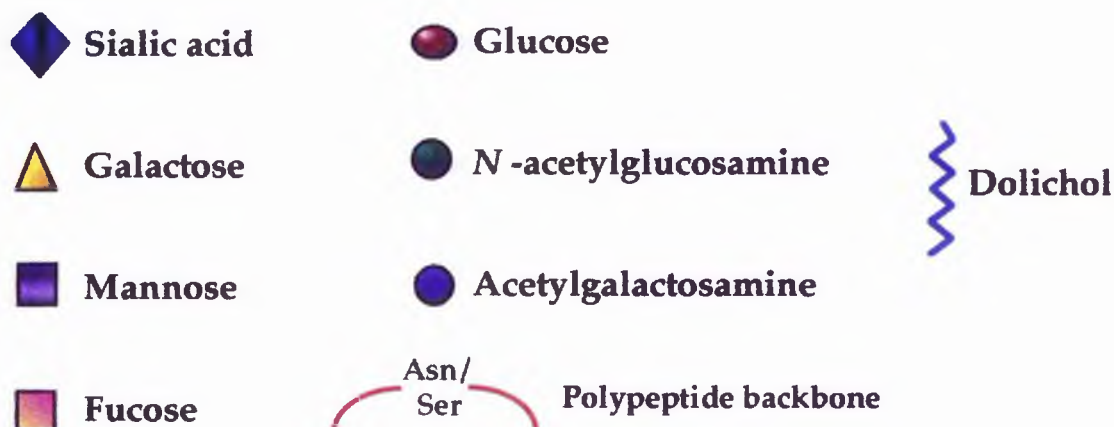
The specification for O-glycosylation is less well understood, but is known to occur at serine or threonine residues and is enhanced by clusters of these amino acids and by the presence of proximal prolines. Addition of O-linked saccharides does not involve a lipid-carrier intermediate and can thus occur in the cytosol, unlike N-glycosylation.

Figure 15. N-glycosylation & O-glycosylation processes occurring within the mammalian-cell exocytotic pathway (adapted from Wolfe, 1993).

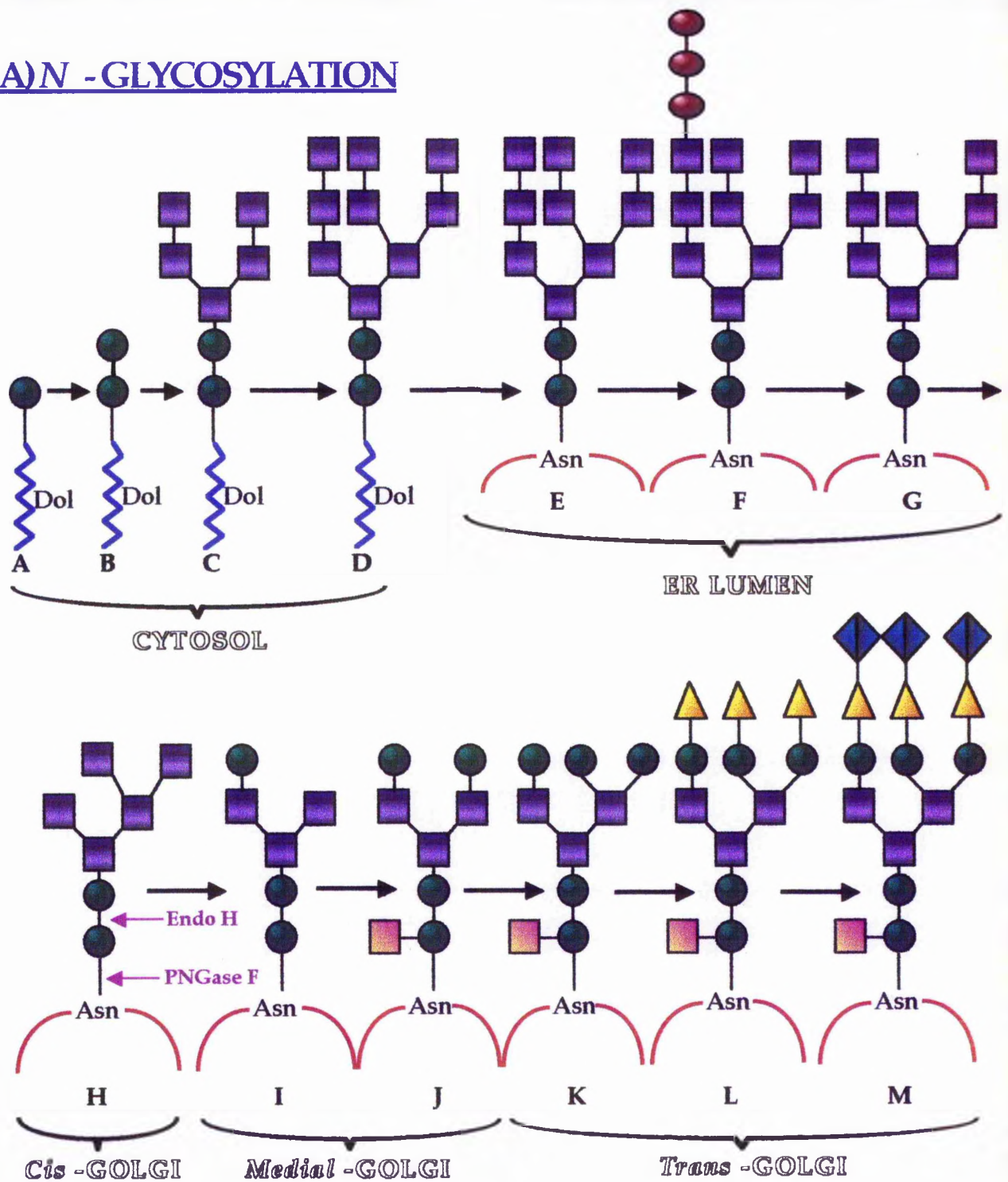
Panel A) The assembly of *N*-linked carbohydrate groups in the ER and golgi complex. Dol, dolichol; Asn, asparagine. The process begins in the cytosolic side of the ER membrane, with transfer of an acetylglucosamine residue from a nucleoside diphosphate carrier to the dolichol unit (A). A further acetylglucosamine residue is added to the first (B). Five mannose residues are then transferred from their nucleotide carriers to the dolichol-sugar complex, forming a branched structure (C). The dolichol-carbohydrate complex is then transported across the ER membrane by the dolichol carrier. In the ER lumen four more mannose units are added (D) with the resulting structure being transferred, as a unit, from dolichol to an asparagine residue in the nascent polypeptide chain (E). In some organisms this step completes the core structure, in others three glucose residues are added via nucleoside carriers (F). With most glycoproteins the three glucose units and one mannose residue are subsequently removed while the protein is still in the ER (G). The glycoprotein is then transported to the Golgi complex, where three mannose residues are usually removed, leaving the core structure shown in (H). Steps (I through M) outline a typical mammalian pathway removing further mannose residues and adding; acetylglucosamine, fucose, galactose and sialic acid to complete the carbohydrate group. Also shown are the cleavage sites for the glycosidases PNGase F and EndoH (PNGase F cleaves most types of carbohydrate chains from glycoproteins. EndoH does not cleave the carbohydrate chains of complex type oligosaccharide additions (Lanzetta *et al.*, 1979)).

Panel B) The assembly of *O*-linked carbohydrate groups in the exocytotic pathway. Ser, serine. A single mannose or acetylgalactosamine residue is transferred from a nucleotide carrier directly to a serine or threonine residue or, more rarely, to a hydroxyproline or hydroxylysine unit. This initial core glycosylation apparently occurs on the lumen side of ER membranes or while the glycoprotein is in transit from the ER to the Golgi complex. To this, or to other simple core structures, additional sugars, including the same variety used in *N*-glycosylation, are transferred from nucleotide carriers within the Golgi complex to complete the oligosaccharide group. The structures shown are one of many possibilities.

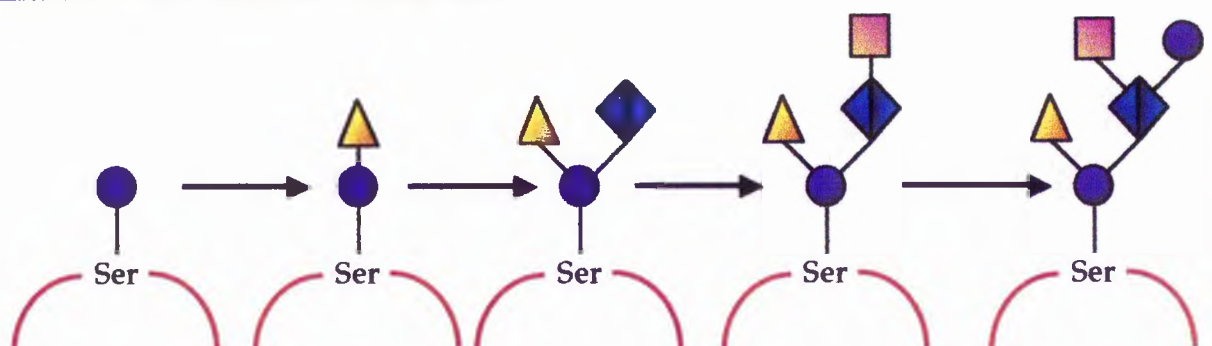
KEY



A) N - GLYCOSYLATION



B) O - GLYCOSYLATION



5 / Retention of proteins within the ER.

Proteins are retained within the ER primarily for two reasons: they are naturally located to the ER, or they are misfolded and are thus destined for degradation. In the latter scenario, misfolded proteins appear to have a stable association with ER chaperone proteins such as GRP78-BiP and do not exit the ER. These misfolded proteins are eventually degraded by, as yet, uncharacterised ER proteases, although possibly involving a 60Kda thiol protease (reviewed in Gaut & Hendershot, 1993). In the former case, proteins are located to the ER by two possible mechanisms; via a retrieval receptor located in the golgi complex, or by inherent properties of the protein preventing forward transport. Of these mechanisms the retrieval mechanism is the best understood.

Proteins containing the sequence lys-asp-glu-leu (KDEL) or related sequences at the C-terminus of the protein are generally located to the ER via retrieval from the CGN by a KDEL-recognising receptor, hERD2, which binds the ligand in a pH-dependant manner. Intriguingly, some bacterial toxins (e.g. *Pseudomonas* exotoxin A, cholera toxin & *E.coli* heat-labile toxins) possess KDEL or related sequences and may gain access to the cytoplasm via the ER (reviewed in Pelham, 1990; Lewis & Pelham, 1992; and Pelham, 1992). Furthermore, studies have shown that at least two other ER localisation signals (dilysine and diarginine) occur, which, like KDEL, act as retrieval signals. Various type I glycoproteins (N-terminus in the lumen) possess a dilysine signal (K(X)KXX, where X is any amino acid) at positions -3 & -4/-5 relative to their C-terminus, which appears to be sufficient and necessary for localisation to the ER. Moreover, with various type II glycoproteins (C-terminus in the lumen) the signal appears to consist of two arginine residues (RR) which are located within the first five amino acids of the N-terminus. Whether a specific receptor is required is not known, but the dilysine motif has been shown to bind to coatamer, a component involved in vesicular transport (Letourneur *et al.* 1994 and Cosson & Letourneur, 1994). KDEL, dilysine and diarginine are not absolutely required for ER localisation since proteins occur which possess these signals and are not localised to the ER. Furthermore, removal of these motifs from some resident ER proteins does not always result in the loss of ER-localisation (reviewed in Nilsson & Warren, 1994). Recently, Mallabiabarrena *et al.* (1995) characterised a possible third

ER retention-signal based upon a tyrosine helix-motif, which was reminiscent of the β -turn structure adopted by tyrosine-containing endocytosis signals.

6 / Transport of proteins from the ER to the Trans golgi network (TGN).

Upon exiting the ER, proteins enter a series of membranous organelles termed the Golgi complex (figure 12). The traditional view of the Golgi complex is of flattened cisternal membranes arranged in polarised stacks termed the; *cis* Golgi network (CGN), *cis* Golgi, *medial* Golgi, *trans* Golgi and the *trans* Golgi network (TGN). Proteins move vectorally through the Golgi complex (see below) from the CGN to the TGN unless they possess appropriate retention signals (see below). The Golgi complex plays a pivotal role in the sorting and post-translational modification of lipids and proteins. Well characterised resident Golgi proteins include the integral membrane glycosyl transferases (*medial* & *trans* Golgi), of which there are an estimated 200 enzymes, which trim and add various sugars to glycoproteins and glycolipids (figure 13). Also present are numerous proteases, such as the furin/PACE protease which is resident in the TGN and is involved in the post-translational cleavage of many cellular and viral glycoproteins (reviewed in Rehemtulla & Kaufman, 1992).

Proteins which are destined only for the Golgi complex possess signals for their appropriate localisation. Golgi localisation signals (GLS) appear not to be as simple, and as well characterised, as those for ER localisation. GLS may be composed of uncharged polar residues (mainly asn, thr & gln) which lie on one face of a predicted alpha-helix within the first transmembrane domain of the protein. This was determined from work carried out with the resident glycosyl transferases and with the infectious bronchitis virus (avian coronavirus) M protein. The localisation signal for resident TGN proteins appears to be located in the cytoplasmic tail, possibly involving tyrosine residues (tyr-gln-arg-leu) as also found on some proteins which are recycled from the cell-surface to the TGN. The mechanism of retention is, as with the signal, not well understood and may be composed of a retrieval receptor analogous to hERD2 for KDEL signals. The other possibility is that of a true retention

mechanism due to the formation of a high molecular weight oligomer or lattice structure, as found with the IBV M protein.

Exactly how proteins travel through the Golgi complex is unclear and currently three models have been hypothesised. (1) Travel by bulk flow, where proteins exit the ER and travel to their destination by massive non-selective vesicular transport. (2) Trapping, where initial transport to the CGN is by bulk flow where, in the CGN, the secretory and cell-surface proteins interact with receptors which flow to the next stage and release their cargo. (3) Receptor-mediated transport, where proteins bind to receptors within the ER and are transported by a small number of vesicles to the CGN (reviewed in Pelham, 1991).

7 / Transport of proteins from the TGN to lysosomes or the cell-surface & external milieu.

When a protein reaches the TGN a variety of options await. Lysosomal proteins (which possess a mannose-6-phosphate targeting signal) are directed to lysosomes, whilst secretory and cell-surface proteins are packaged into secretory vesicles for transport to the plasmamembrane or the external milieu. This latter pathway of transport is constitutive to most cell-types. However, specialised secretory cells such as those found in exocrine, endocrine and neuronal tissue can, in addition to constitutive secretion, regulate secretion through the formation of dense secretory granules which are released upon receiving an external stimulus. A sorting mechanism for the different pathways has not been clearly elucidated yet (reviewed in Chidgey, 1993).

CHAPTER 5

OBJECTIVES OF THE PROJECT

Epitope-tagging of proteins has recently gained widespread use for the detection and affinity-purification of recombinantly expressed proteins, and involves the addition of a set of amino acid residues, via recombinant DNA technology, to either terminus, both termini or within the protein sequence of choice. Oligopeptides which have been utilised for the labelling of proteins include; HA-tag (derived from the haemagglutinin of influenza virus), Myc-tag (derived from the myc oncogene), BTag (derived from the major core protein, VP7, of bluetongue virus (Wang *et al.*, 1996)), flag-tag and His-tag (linear sequence of variable length, but usually 6 histidine residues). These tags possess different characteristics to each other in terms of length, charge and affinity of interaction to their respective mAbs (or chelating agent), and therefore there is no universal epitope tag which can be employed to successfully label every type of protein, expressed in any expression system.

A further addition to the epitope-tagging repertoire is a 14 amino acid oligopeptide, termed Pk-tag, which was derived from the P/V proteins of the paramyxovirus SV5. The Pk-tag has previously been shown to retain its antigenicity to a mAb (termed anti-Pk) when attached to the N or C-terminus of proteins expressed in bacteria or baculovirus infected insect cells (Hanke *et al.* 1992, 1994a & 1995). Furthermore, Pk-tag & anti-Pk mAb have successfully been utilised for the detection of proteins expressed in yeast cells as part of the yeast-two-hybrid system (R.E. Randall, personal communication).

Furthermore, the addition of the Pk-tag to proteins destined for secretion from bacteria, did not interfere with the secretion process (T. Hanke (unpublished observations)

and E.A. Green (in press)). Similarly, the presence of the Pk-tag on the C-terminus of the envelope glycoprotein from simian immunodeficiency virus (SIV) gp160, was shown not to alter the conformation, binding to CD4, or antigenicity of the Pk-tagged gp160, when expressed in an insect-cell system and compared to the parental gp160 (Hanke *et al.*, 1995). Further beneficial features of the Pk-tag epitope are the weak immunogenicity of the peptide (Randall *et al.*, 1993), and the lack of cross-reactivity to host-cell proteins in the systems tested. Furthermore, in the context of multivalent recombinant subunit vaccine development it has been demonstrated that the Pk-tag and anti-Pk mAbs are useful reagents for the construction of SMAA complexes as a delivery system for multiple immunogens.

An extensive analysis of the processing of Pk-tagged glycoproteins when produced in mammalian cells has not been carried out. For many purposes, including recombinant subunit vaccine development, proteins may have to be produced in mammalian cells to ensure correct processing and post-translational modification. Moreover, it has been demonstrated that the addition or deletion of amino acids at the N or C-terminus of mammalian-cell membrane-bound glycoproteins may affect their processing through the exocytotic pathway, resulting in retarded transport, aberrant glycosylation and inhibition of oligomer formation (Parks & Lamb, 1990; Simpson & Lamb, 1992 and Sergel & Morrison, 1995), all of which may alter the conformation of the protein relative to the wildtype protein.

Therefore, an investigation was carried out to determine whether the addition of the 14 amino acid Pk-tag to the N or C terminus of a model type II glycoprotein (and of truncated secretory forms), and to the C-terminus of a model type I glycoprotein (and of a truncated secretory form), would result in aberrant processing and antigenicity of the protein. The model proteins used in this thesis were the well characterised envelope glycoproteins derived from the paramyxovirus SV5: haemagglutinin-neuraminidase (type II topology) and Fusion (type I topology).

Initially the Pk-tagged proteins would be transiently expressed in mammalian cells via the vaccinia virus/T7-RDRP system with subsequent analysis of exocytotic processing and antigenicity. In addition, if the Pk-tagged glycoproteins were successfully expressed in a transient system, then the construction of expressing cell-lines, via the construction of EBV-

based vectors, would be attempted so as to produce enough protein for immunogenicity and protection trials.

Furthermore, conventional methods for producing large quantities of recombinant protein in mammalian cells have relied upon the amplification of the gene copy number when integrated into the cellular genome. Such methods are extremely time-consuming and employ large quantities of toxic compounds. Therefore, the construction of a novel inducible system, based-upon the SV40 LTA_g and SV40 *ori*-containing EBV-based plasmids, in which episomal plasmids can be transiently amplified for a short-term increase in recombinant protein production, was attempted.

MATERIALS AND METHODS

1/ Bacterial strains.

E. coli K12 strains; XL-1 blues (in house supply) [genotype->*rec A1, end A1, gyr A96, thi -1, hsd R17, sup E44, rel A1, lac* , {F⁺ *pro AB, lac IqZΔM15, Tn10 (tet^r)*}] or Top 10s (Invitrogen Corporation) [genotype->*mcr A, Δ(mrr-hsd RMS-mcr BC), ø80Δlac ΔM15, Δlac X74, deo R, rec A1, end A1, ara D139, Δ(ara, , leu) 7679, gal U, gal K, rps L(Str^r), nup G*] were used for routine propagation of plasmids. Bacteria were grown in Luria-Bertani (LB) broth with antibiotics added when required.

2/ Mammalian tissue culture cells and antibodies.

293s (transformed human embryonic kidney cells) at low passage number (initially no. 28) for efficient transfection were a kind gift from F. Graham (McMaster University). 293 cells were cultured in Glasgow's` Modified Essential Medium (GMEM) supplemented with 10% Foetal Calf Serum (GMEM-10) and passaged weekly. Cos-7 cells (transformed African Green- monkey kidney cells) were generously donated by F. Arenzana-Seisdedos (Institute Pasteur) and passaged twice weekly in GMEM -10. Hela cells (human cervical carcinoma cells) were passaged twice weekly in GMEM-10. The 293 cell-line that constitutively expresses the Epstein-Barr virus nuclear antigen 1 (EBNA-1) gene product was purchased from Invitrogen Corporation and maintained as described for parental 293s. Monoclonal antibody SV5-P-k recognising an epitope on SV5 phospho (P) and V proteins and monoclonal antibody K3 recognising dinitrophenol were isolated (mAb SV5-P-k, Randall *et al.*, 1987) and their specificity's were determined (mAb SV5-P-k, Southern *et al.*, 1991) in our laboratory.

Monoclonal antibody to bacterial β -galactosidase was purchased from Boehringer Mannheim. Conformationally-dependant mAbs to SV5 HN and F were isolated and characterised in our laboratory (Randall *et al* 1987). MAb (PAB 416) to SV40 LTA_g was generously donated by D. Lane (University of Dundee, Scotland).

3/ Plasmids.

Plasmids pGEM-3zf(+) and pGEM-9zf(-) were purchased from Promega Corporation. The EBV mammalian expression vectors; pCEP4, pREP4 & pREP8 and the transient mammalian expression vectors pcDNA1/AMP & pcDNA3 were obtained from R&D Systems Europe Ltd. The tetracycline-responsive mammalian expression vectors pUHD10-13 (tTa-driver), pUHD10-15 (β -gal. reporter) and pUHD13-3 (Luciferase reporter) were a kind gift from A.G.H. Bujard. Plasmids containing the cDNA of the HN protein of SV5 and the cDNA of the SV5 Fusion protein were generously donated by R.A. Lamb (Evanston, Illinois). Plasmid pPE346 which contains the β -galactosidase gene under control of the hCMV IE enhancer/promoter was kindly supplied by N. Stow (MRC Virology, Glasgow). Plasmid pMR101 containing the Foot and Mouth Disease virus 2A cleavage sequence adjoined to the aminoglycoside phosphotransferase gene was obtained from M. Ryan. Plasmid pSVori encoding the SV40 LTA_g was obtained from J. Doherty (Bute Medical Building, University of St. Andrews). Plasmid pMHVA encoding the Adenovirus VAI I gene was donated by R.E. Randall (University of St. Andrews). Plasmids pEBV-tTa/neo and pMRtTa/neo were generously donated by B. Precious (St.Andrews Uni). Plasmid pCMVluc was generously donated by J.F. Wright (Univ. of St. Andrews).

4/ Agarose-gel electrophoresis.

DNA samples were separated via horizontal agarose-gel electrophoresis using a mini-gel (10cm x 10cm) apparatus (Bioscience 101). T.B.E. or T.A.E. buffer was used depending on the particular application. The agarose (molecular biology grade, Sigma) concentration ranged from 0.8% to 2% depending on the size of the DNA fragments to be separated. Gels

were ran at varying voltages and for differing times, again depending on the application. The DNA was subsequently analysed by ethidium bromide staining and visualisation under U.V.

5/ Extraction of DNA fragments from agarose-gels.

DNA fragments were separated on an agarose-gel prepared from TAE buffer. Gel fragments containing the DNA band of interest were excised from the agarose gel and processed using Qiaquick gel-extraction spin-columns (Hybaid Ltd.) according to the manufacturers recommendations. To assess recovery of the DNA fragments, $1/10$ th aliquots (5µls) were visualised after electrophoresis through agarose-gels and staining with ethidium bromide.

6/ DNA linkers and primers.

Oligonucleotides used for PCR amplification of SV5 genes were synthesised in-house using an Applied Biosystems oligonucleotide synthesiser. Oligonucleotides were purified via ethanol precipitation as described in Manniatis *etal*. Oligonucleotides used for the generation of Pk-tag-linkers were synthesised and H.P.L.C. purified by Oswell's DNA service (Chemistry department, University of Edinburgh, now located at the University of Southampton).

Oligonucleotides of positive and negative sense for the generation of the C-terminal Pk-tag linker were -> **GEM ctPkpos** 5'-CTG GGC GCG CCA ATA ATC TTA AGG GAA AGC CGA TCC CAA ACC CTT TGC TGG GAT TGG ACT CCA CCT AAC -'3 and **GEMctPkneg** 5'-TCG AGT TAG GTG GAG TCC AAT CCC AGC AAA GGG TTT GGG ATC GGC TTT CCC TTA AGA TTA TTG GCG CGC CCA G - '3. The two oligonucleotides were annealed together and 5' phosphorylated using standard protocols (Manniatis *etal*). Oligonucleotides of positive and negative sense for the generation of the N-terminal Pk-tag linker were -> **GEM ntPkpos** 5'-CCG GGA CCA TGG GAA AGC CGA TCC CAA ACC CTT TGC TGG GAT TGG ACT CCA CCT GGC GCG CCT AGA ACC TTA AGT AAT -'3 and **GEMntPkneg** 5'-CTA GAT TAC TTA AGG TTC TAG GCG CGC CAG GTG GAG TCC AAT CCC AGC AAA GGG TTT GGG ATC GGC TTT CCC ATG GTC -'3. The two oligonucleotides were

processed as for the ctPk oligonucleotides. Underlined sequences correspond to the Pk-oligonucleotide sequence.

Primers annealing to the non-coding and coding strands of SV5 genes, designated **FOR** (forward) and **BACK** (backward), respectively, were as follows; **hn, FOR** 5'- ATT ATA AGG CGC GCC ATG GTT GCA GAA GAT GCC CCT G -'3, **hn, BACK** 5'- TTA AAT ACT TAA GGG ATA GTG TCA CCT GAC GGA T -'3, **f, FOR** 5'-ATT ATA AGG CGC GCC ATG GGT ACT ATA ATT CAA TTT CTG -'3, **f, BACK** 5'-TTA AAT ACT TAA GTT TAT GAT AAA CAA AAT TCT C -'3. The primer used for the generation of the secretable form of the fusion protein was, **sf, BACK** 5'-TTA AAT ACT TAA GGT CAC TTT GTG CTA AGT GTT G -'3. The primer used for the secretable form of the HN protein was, **sHN, FOR** 5'- ATC TGG CGC GCC GAG AGT TTA ATA ACC CAA AAG -'3. The primer used in the repair of mutated Pk-tag sequence in pT7ctPk/ss was, **MR, BACK** 5'- ACT CAA GCT TGC ATG CCT GCA GGT CGA G -'3. The primer used for the generation of native HN protein was, **hn(taa), BACK** 5'- TTA AAT CTT AAG TTA GGA TAG TGT CAC CTG ACG GAT -'3.

The Primers used for the amplification of the hygromycin B phosphotrans-ferase gene and thymidine kinase poly A sequence were, **hygB, FOR** 5'-ATT AGG GCC CAT GAA AAA GCC TGA ACT CAC C-'3, and **TKpA, BACK** 5'- ATT AGG GCC CTC AGT CCA GTC GTG GAC CAG A. The primers used for the amplification of the SV40 virus large T-antigen were as follows; **LTA_g, FOR** 5'- ATA TGA ATT CAC CAT GGA TAA AGT TTT AAA CAG AG-'3, and **LTA_g, BACK** 5'- TAT ATG CAT GCT TAT GTT TCA GGT TCA GGG GG-'3. The positive and negative sense 5'-phosphorylated oligonucleotide linkers used for the introduction of a second Xho1 restriction endonuclease site in pTETHNntPk and pTETHNctPk were; **Xho1, pos** 5'-AGC TTA CTC GAG GA-'3 and **Xho1, neg** 5'- ATG AGC TCC TTC GA-'3. The two oligonucleotides were annealed together using standard protocols (Manniatis *etal*). 5'-phosphorylated oligonucleotides of positive and negative sense used for the construction of an optimal Kozak sequence linker were; **Kozak, pos** 5'- GAA TTC GCC GCC AC-'3 and, **Kozak, neg** 5'- GTG GCG GCG AAT TC-'3. The two oligonucleotides were annealed together using standard protocols (Manniatis *etal*).

7/ Polymerase Chain Reaction (PCR).

Amplification of cDNAs was carried out using P.C.R. methodology. The 100 μ l reaction mixture consisted of 0.2mM of each dNTP, approximately 1ng of template DNA, 1 μ M of each primer and PCR buffer (x1: 10mM Tris-Cl pH8.8, 1.5mM MgCl₂, 50mM KCl and 0.1% Triton X-100). Dynazyme I or Dynazyme II DNA polymerases (Finnzymes Ltd.) were the last components added to the mixture, which was then overlain with 50 μ l of liquid paraffin to minimise evaporation. Typically, 30 cycles of; 94°C for 1.5 mins., 55°C for 1.5 mins. and 72°C for 2 mins. with the last cycle being 72°C for 9.9 mins.

8/ Processing of Amplified cDNAs for Ligation.

The liquid paraffin was pipetted from the reaction mixture and subsequent removal of polymerase, primers and salts was achieved by utilising PCR-purification spin columns (Flowgen Ltd.) which were used according to the manufacturers recommendations. Purified DNA was digested with the appropriate restriction endonucleases for 10-16 hrs according to the vendors instructions. The digested products were agarose-gel-purified (above) and used for subsequent ligations.

9/ DNA Ligations.

Depending on the size of DNA fragments to be ligated, an approximate 3:1 (large DNA inserts, >500bp) to 15:1 (small DNA inserts, <100bp) molar ratio of insert to vector fragments were combined in a total volume of 10-20 μ l of ligation buffer (x1: 50mM Tris-Cl pH7.8, 10mM MgCl₂, 10mM DTT, 1mM ATP and 25mgml⁻¹ BSA) where-upon 200-1000U of T4 DNA ligase (New England Biolabs) were added and the ligation reaction incubated at 16°C for 10-16 hrs. Half of the ligation reaction was used for subsequent transformation of competent bacteria.

10/ Preparation of competent bacteria.

E. coli strains were made chemically-competent using MgCl₂ and CaCl₂ as follows. A 100ml bacterial culture was grown to an OD₆₀₀ = 0.4-0.6, cooled on ice and pelleted at

x3,000g for 10 mins at 4°C. The pellet was gently resuspended in 10ml of ice-cold, sterile, 0.1M MgCl₂ and left on ice for approximately 3mins. The suspension was spun as above and the pellet resuspended in 2mls of ice-cold, sterile, 0.1M CaCl₂. The bacteria were optimally competent for transformation 12-24hrs after preparation.

11/ Transformation of competent bacteria.

200µls of competent cells were incubated with 10µls of ligation reaction, on ice, for 30 mins. with periodic mixing end-over-end. 1ng of pUC19 plasmid DNA or no DNA were added to competent cells as positive and negative controls respectively (typically 10⁶ to 10⁷ transformants per µg of DNA was observed). After incubation, the bacteria were heat-shocked for 1.5 mins. at 42°C and returned to ice for 5 mins. The transformation mix was added to 0.8mls of LB- broth and incubated at 37°C for 1hr, to allow expression of β-lactamase for ampicillin selection, after which 0.2ml of the transformed bacteria were plated on LB-agar containing relevant antibiotics.

12/ Preparation of plasmid DNA.

Small-scale preparations of DNA, mini- preps., were prepared using the alkaline-lysis method of Birnboim and Doly (1979). Large-scale preparations, maxi- preps., were prepared either by utilising Qiagen DNA maxi-prep. ion- exchange columns (Hybaid Ltd.) as used according to the manufacturer, or by a modified large scale alkaline-lysis procedure, as follows. Briefly, a 500ml O/N culture was centrifuged at x10,000g and the pellet resuspended in 4mls of TE buffer (10mM Tris-Cl pH8.0 & 1mM EDTA pH8.0). Eight mls of lysis solution (0.2M NaOH & 1% SDS) were added and the viscous suspension was mixed end- over- end for no longer than 5 mins.. Six mls of ice-cold neutralisation buffer (3M potassium ethanoate pH5.5) were added and the mixture briefly mixed end-over-end and incubated in ice for 15 mins. The precipitate was pelleted at x23,000g for 30 mins. at 4°C. where-upon the plasmid-containing supernatant was precipitated with 1 volume of iso-propanol and pelleted as above. The pellet was resuspended in 2 mls 10mM Tris-Cl pH8.0, added to 2.5 mls of 4.4M

LiCl and incubated on ice for 15 mins. to precipitate out high molecular weight RNA. After centrifuging as above, the supernatant was precipitated with 2.5 volumes of ethanol, pelleted and resuspended in 0.7 mls 10mM Tris-Cl pH8.0. Seven microlitres of DNase-free RNase (Boehringer Mannheim) were added and incubated for 30-60 mins. at 37°C. An equal volume of PEG solution (1.6M NaCl, 16% polyethylene glycol 6000) was added and the mixture vortexed then incubated at 4°C for 10 mins. after which centrifugation was carried out at 4°C, 14,000rpm in a benchtop rotor. The pellet was resuspended in 0.4mls of TE and the plasmid preparation was phenol/chloroform- extracted x3, chloroform- extracted x2 and ether- extracted x1, precipitated with 1/10 vol. NaAc/2.5 volumes of ethanol, pelleted at 4°C, 14,000 rpm in a benchtop microfuge then resuspended in 200-500µls of 10mM Tris-Cl pH8.0. Quality and quantity of DNA was analysed by spectrophotometric readings at 260nm & 280nm and by visualisation on an agarose-gel (Manniat *et al.*).

13/ DNA Sequencing.

DNA sequencing was performed on double-stranded plasmid DNA templates using the dideoxy chain- termination procedure. Plasmid DNA for sequencing was prepared using Qiagen plasmid preparation columns (Hybaid Ltd.), as above. 5µgs of plasmid DNA per primer-reaction were sequenced using the Sequenase® version 2.0 sequencing kit (Amersham International), [$\alpha^{35}\text{S}$]dATP (1×10^3 Ci.mmol⁻¹; Amersham International) and used exactly according to the vendor's recommendations.

14/ Preparation of cationic liposomes for DNA transfection.

The following procedure was kindly given by R. Elliott (MRC Virology, Glasgow) and results in a lipid mixture equivalent to Gibco-BRL's TransfectACE. Briefly, 1ml DOPE (dioleoyl L- α -phosphatidyl ethanolamine, Sigma) was pipetted into a glass universal whereupon 4mgs of DDAB (dimethyldioctadecyl ammonium bromide, Sigma), w/w ratio of 1:2.5 DDAB:DOPE, were added and briefly vortexed till dissolved. The chloroform was evaporated using a stream of nitrogen whilst rotating the universal to ensure an even film of lipid dispersed around the walls. The dried lipids were resuspended in 10mls of sterile

distilled water by a 10-15 min. sonication in a sonicating waterbath. Once resuspended, the turbid mixture was sonicated with a soniprobe (15 microns amplitude) using 30 second pulses and incubating on ice in between, until, or as close to, clear (approx. 20-40 mins.). Lipids were tested against previous batches for transfection efficiency using a β -galactosidase expressing plasmid, pPE346, then stored in 1ml aliquots at 4°C.

15/ Staining protocol for detection of β -galactosidase expressing cells.

Transfected cells were washed once with PBS then fixed for 5 minutes at room temperature by the addition of 1ml of fixative (5% formaldehyde, 2% sucrose in PBS) per 3.5cm diameter dish. The fixative was removed and the cells washed with PBS. Three millilitres of freshly prepared staining solution (5mM potassium ferricyanide, 5mM potassium ferrocyanide, 2mM MgCl₂ and 0.5mgml⁻¹ X-gal, in PBS) was added to the washed cells which were then incubated at 37°C overnight.

16/ Quantitative luciferase assays.

Cells grown in 6-well plates were transiently transfected with appropriate luciferase reporter vectors, as detailed above. Thirty six hours post-transfection the cells were washed twice with PBS and lysed with 200 μ ls of luciferase lysis buffer (25mM Tris-phosphate pH7.8, 8mM MgCl₂, 1mM DTT, 1% Triton X-100 and 15% glycerol). The lysate was scraped off the culture dish and the cellular debris was pelleted by a brief 10 second centrifugation. 100 μ ls of soluble cell extract was assayed for luciferase activity in a Lumat LB9501 luminometer (Berthold, Wildbad, Germany) using the integral mode (10 seconds). Luciferin buffer was: 1mM ATP, 0.25mM D-luciferin (SIGMA), 1%BSA in luciferase lysis buffer. Luciferase activity is expressed as relative light units (RLUs) per μ g of protein. Protein determination was performed using the Bradford assay.

17/ Production of Vaccinia Virus vTF7-3 stocks.

For the production of small stocks of virus, a recently confluent Hela monolayer (25cm² flat) was infected with virus (in house supply) at a m.o.i. of 1, in GMEM

supplemented with 1% FCS (GMEM-1) for 1-2 hrs, where-upon GMEM-10 was added and the cells incubated at 37°C for 48 hrs. The cells were scraped into the medium, pelleted and resuspended in 2mls 10mM Tris-Cl, 1mM EDTA, pH 7.0 and the virus released by three cycles of freeze-thawing, and stored at -70°C. Medium stock of virus were prepared as above with the whole of the small virus stock being used to infect $\times 10^7$ 75cm² flats of recently confluent Hela cells. The titre of the midi-prep. was determined (see below).

Large stocks of semi-purified virus was prepared as follows; 1L of Hela S3 spinner cells (5×10^5 cells ml⁻¹) were concentrated to a density of 5×10^6 cells ml⁻¹ by low-speed centrifugation and resuspension in GMEM-1. Virus at a m.o.i. of 5 (previously treated with 0.1 vol. of trypsin (2.5 mgml⁻¹, Sigma) for 30 mins. at 37°C) was added to the concentrated cells and incubated, with constant stirring, at 37°C for 1-2 hrs, where-upon 900mls of GMEM-10 were added and incubation was continued for a further 48 hrs. Infected cells were pelleted and resuspended in 20mls of 10mM Tris-Cl pH 9.0 at 4°C (all subsequent steps were carried out on ice). The suspensions were homogenised with 30 strokes of a dounce homogeniser and the nuclei pelleted by centrifugation at 750xg for 5 mins. The supernatant was decanted and the pellet resuspended in 20mls 10mM Tris-Cl pH 9.0 and homogenised and pelleted as above. The supernatants were combined, 0.1 vol. of trypsin was added and incubated as above. The trypsin treated virus stock was layered onto an equal volume of 36% sucrose in 10mM Tris-Cl pH 9.0 in a Beckman SW28 swingout tube and centrifuged at 13,500 rpm (25,000xg) for 80 mins. at 4°C. The pellet was resuspended in 2mls of 1mM Tris-Cl pH 9.0 and stored in aliquots at -70°C.

18/ Titration of Vaccinia virus stocks.

Duplicate 35 mm dishes of 293 cells which had just reached confluence were washed with GMEM-1 and infected with 10-fold dilutions of virus in GMEM-1 for 1-2 hrs at 37°C. The viral inoculum was removed, GMEM-10 was added, and the infected cells incubated for a further 48 hrs at 37°C, where-upon the medium was removed and the monolayer stained with a solution of 0.1% Crystal Violet in 2% ethanol for 2-3 mins. at room temperature to visualise the areas of cell degeneration (plaques).

19/ Transient protein production.

a) Expression from T7 ϕ 10 promoter.

Mammalian cells were passed into 9cm diameter petri dishes or individual wells (3.5cm diameter) of a 6-well dish in numbers sufficient to be confluent the next day. Confluent monolayers were washed once in Optimem- 1 (Gibco-BRL) and recombinant vaccinia virus which expresses T7 RNA polymerase, diluted in Optimem- 1 to give a m.o.i. of 2, was added to the monolayers and incubated at 37°C, with periodic agitation, for 1hr. Plasmid DNA was introduced into the infected cells using a liposome-mediated procedure. Briefly, for one 3.5 cm diameter well, three micrograms of DNA were added to 100 μ ls of Optimem-1 and mixed. 10 μ ls of the liposome mixture were added to the 100 μ ls of Optimem-1/DNA, gently mixed and left to stand for 10 - 15 mins. at room temperature. All steps were carried out in polystyrene tubes or glass bijoux due to the affinity of the liposomes for polypropylene. The liposome/DNA mixture was carefully added to the infected cells which were incubated for a further 12-18 hrs.

b) Expression from the human CMV I.E. enhancer/promoter in 293 cells.

Dishes were seeded with cells to obtain 40-50% confluency the next day. Liposome/DNA mixtures were prepared as above except that after the 10 -15 minute incubation, the volume was adjusted to 1ml which was directly added to cells which had been washed once with Optimem-1. The Optimem-1/liposome/ DNA mixture was left on the cells for 12-16 hrs after which the cells were washed once with GMEM-10 and incubated for a further 24-72 hrs in GMEM-10.

c) Expression from the human CMV I.E. enhancer/promoter in Cos-7 cells.

Dishes were seeded with cells to obtain 40-50% confluency the next day. Transfections were carried out using DEAE-Dextran methodology. Briefly, for each 35mm dish, 1.0 μ g of DNA were added to 500 μ ls Optimem-1 where-upon 50 μ ls of a 1mgml⁻¹ filter-sterilised DEAE-Dextran (5x10⁵ mwt., Sigma) solution (prepared in TBS) were added and the mixture mixed. The DNA/DEAE-Dextran mixture was incubated for 15 mins. at room temperature. Medium from the cells was aspirated and the cells washed twice with PBS. The DNA/DEAE-

Dextran mixture was brought up to 1ml with Optimem-1 and the mixture added to the cells for 1 hr, 37°C, 5% CO₂ with periodic rocking. The transfection mixture was removed and the cells washed twice with PBS where-upon 3mls of GMEM-10 containing 100µM chloroquine (Sigma) were added and the cells incubated for 4 hrs at 37°C, 5% CO₂. The medium was removed and the cells were washed once with GMEM-10 and incubated for a further 48-72hrs in GMEM-10.

d) Expression from the tetracycline-responsive promoter in 293 cells.

Transfections were carried out as for vectors containing the hCMV IE enhancer/promoter except that a 1:1 ratio of driver to responder plasmids were mixed together to give a final amount of 3µgs which was used to transfect 1x10⁶ cells, in the absence or presence of tetracycline.

20/ Synchronisation of Mammalian cells at G2/M phase.

Cells were seeded at 40-50% confluency in tissue culture vessels. Synchronisation at the G1/S phase of the cell cycle was achieved by an overnight (usually 16-18 hrs) treatment with 3.3mM thymidine (Sigma, prepared as a 0.2M stock in water, filter sterilised and stored at -20°C). The cells were released from the thymidine blockade and set-up for G2/M synchronisation by washing the cells in growth medium followed by an overnight incubation in growth medium containing 0.1µgml⁻¹ Hoechst 33342 (Sigma, prepared as a 0.1mgml⁻¹ stock in DMSO, stored at -20°C in a light-proof eppendorf).

21/ Production of stable cell lines.

A 40% confluent 25 cm² flat of tissue culture cells, synchronised at the G2/M phase of the cell-cycle, was transfected with the required amount of DNA, as above. Twelve hours after transfection the cells were washed with GMEM-10 and incubated for approximately 24 hrs in GMEM-10 to revive from the transfection. The cells were then seeded into 12cm diameter petri dishes, where-upon selective agent was added (hygromycin B (Boehringer Mannheim), Histidinol dihydrochloride (Sigma or Fluka) and Geneticin (Gibco BRL) were used at working concentrations of 200µgml⁻¹, 1mM to 5mM, and 400µgml⁻¹ respectively. The

cells were incubated in selective medium, replaced every 3 days with fresh selective medium, until individual colonies grew-up which were then isolated by ring-cloning. Isolated colonies were screened for expression by immunofluorescence, Western Blot and immunoprecipitation analysis. For the production of stable clones expressing from the tetracycline-responsive promoter, tet^C at a working concentration of 3µgml⁻¹ was included in all the above steps and removed 24-48hrs prior to screening by washing the cells x4 with growth medium.

22/ Determination of protein concentration.

Protein concentrations were determined using either the method of Bradford (if DTT or β-mercaptoethanol was present in the sample) or the BCA protein assay (Pierce Ltd.). For the Bradfords assay 10µls of sample was mixed with 990µls of Bradford's reagent (100mg of coomassie blue G250, 100ml of orthophosphoric acid, and 50ml of absolute ethanol, made up to 1L with distilled water) and incubated at room temperature for 5 minutes followed by measuring the absorbance at 595nm on an LKB Biochrom Ultraspec II spectrophotometer, with concentrations been determined from a BSA standard curve. The BCA assay was performed using the microtiter plate method as described in the manufacturers instructions, again with BSA as the standard curve.

23/ SDS-PAGE and Western Blot Analysis.

Protein samples were suspended in disruption buffer (x1; 2% w/v SDS, 5% v/v β-mercaptoethanol, 2.5% v/v glycerol and 0.25 % w/v bromophenol blue) boiled for 5 mins. and separated through 6-12% SDS-polyacrylamide gels using thin (0.75 mm) mini-slab gels of the Bio-Rad mini-protean II electrophoresis system. Separated polypeptides were transferred (transfer buffer- 48mM Tris, 39mM glycine, 20%v/v methanol and 0.0375% SDS) onto nitrocellulose or P.V.D.F. filters using a semidry gel electroblotter (CBS Scientific Co.). The filters were blocked with PBS containing 5% (w/v) skimmed milk powder and 0.1% (v/v) tween 20 then incubated with monoclonal or polyclonal antibodies diluted 1:200 to 1:1000 in blocking buffer. After washing in blocking buffer, bound antibodies were detected using HRP-conjugated protein A (Amersham International) or HRP-conjugated anti-mouse

antibodies (Amersham International) diluted in blocking buffer, washed as above, then developed by the enhanced chemiluminiscence (ECL; Amersham International) protocol according to the manufacturers recommendations.

24/ Stripping of Nitrocellulose or P.V.D.F. membranes.

Membranes were incubated in stripping buffer (0.15ml β -mercaptoethanol, 2ml 20% SDS, 1.25ml 1M Tris pH6.7, 16.6ml H₂O) in a hybridisation oven at 72°C for 30 minutes. The membrane was washed once for 5 mins, then for 15 mins three times, in 250ml PBS containing 0.1% (v/v) Tween 20. The membrane was subsequently blocked for 1 hour in PBS containing 5% skimmed milk powder and 0.1% (v/v) Tween 20 and probed with anti-sera as detailed above.

25/ Immunoprecipitation.

Soluble antigen extracts (SAE) were prepared by washing transfected cells with ice-cold PBS followed by lysis using ice-cold immunoprecipitation buffer (IPB, 0.1% SDS, 0.5% NP40, 0.1% NaN₃, 0.65M NaCl, 1mM EDTA and 20mM Tris-Cl pH 7.5) and a 15 min. incubation on ice. The viscous cellular debris was scraped off the dish, briefly sonicated and insoluble material pelleted by ultra-centrifugation at 400Kg, at 4°C. Soluble antigen extracts were incubated with 1 μ l of ascites fluid per 200 μ ls of SAE for 1-2 hrs on ice followed by the addition of 20 μ ls of a 10% killed and fixed suspension of the Cowan A strain of *S. aureus* for 1-2 hrs on ice. The precipitates were pelleted by a 3 min. centrifugation at 6,500 rpm in a bench top microfuge and the pellet washed x3 by the addition of IPB and subsequent pelleting. The washed pellet was resuspended in 20-50 μ ls of x1.5 SDS-disruption buffer and processed as above for SDS-PAGE/W. blot.

26/ Immunofluoresence.

Cells to be probed were grown on 10mm or 13mm glass coverslips in individual wells of a 6 well or 24 well tissue culture plate. Cells to be fixed were washed twice with PBS followed by a 10 minute incubation in fixative (5% formaldehyde, 2% sucrose in PBS). The

fixative was aspirated and the cells were washed twice in PBS followed by a 5 min. incubation in permeabilisation buffer (0.5% NP40, 10% sucrose, 1% calf serum(CS) in PBS) if required. The cells were washed x3, in PBS/1% CS. 10-20µls of a 1:100 dilution of primary antibody was spotted onto siliconised glass microscope slides and the fixed coverslips were placed onto the antibodies and incubated for 30 mins. at 37°C in a humidified chamber. The coverslips were washed by sequential multiple dipping in permeabilisation buffer x1, PBS/1% CS x3, then PBS x2. The coverslips were then placed onto 10-20µls of 1:100 diluted Texas Red-conjugated second antibody and incubated for 30 mins. at 37°C in the dark, again in a humidified chamber. The coverslips were washed as previous and mounted on glass microscope slides using 2.5µls of Citifluor Af-1 mountant.

27/ Colony Hybridisation.

A 9cm diameter nylon membrane (Hybond-N, Amersham International) was overlain onto an agar plate containing the colonies arising from the transformation of a DNA ligation reaction into competent *E. coli* . The membrane was left to moisten for approximately 3 mins., carefully peeled off the plate and placed, for 2 mins., colony-side up onto a 3MM Whatman filter pre-soaked in lysis solution (10% SDS). The membrane was then transferred for 2 mins. colony-side up, onto a 3MM Whatman filter pre-soaked in denaturing buffer (0.5M NaOH, 1.5M NaCl). Next, the membrane was again transferred colony side-up onto a 3MM filter pre-soaked in neutralisation buffer (1M Tris-HCl, 1.5M NaCl) for 5 mins., this step being repeated once more. After washing the membrane briefly in x2 S.S.C. the denatured DNA was cross-linked to the nylon membrane by illumination with ultraviolet light for 3 mins. The filter was covered in prehybridisation buffer (x6 SSC, x5 Denhardts, 0.1% SDS and 50µgml⁻¹ ssDNA) and incubated at 60°C for 4hrs , the pre-hybridisation buffer was decanted and hybridisation buffer (as for prehybridisation buffer but also containing radiolabelled probe) added. Incubation was continued at 60°C overnight. The filter was washed as follows; 6x SSC, 0.1% SDS for 30 mins. at 60°C, 2xSSC, 0.1% SDS for 30 mins. at 60°C, 0.2x SSC, 0.1% SDS for 30 mins. at 60°C, last step repeated then a final brief rinse in 2x SSC before exposure to X-ray film.

28/ Chemical Cross-linking Reactions.

Chemical cross-linking was performed by a method similar to that described by Paterson and Lamb 1993. 293 cells were infected with Vaccinia virus and transfected with the appropriate plasmid as described above. Twelve hours post-transfection the cells were washed with PBS and removed from the dish using the EDTA solution: Versene. The cells were pelleted by centrifugation at 6,500 rpm, for 1 minute, in a benchtop microfuge. The pellet was resuspended in 70µl of PBS and the cells solubilised by addition of NP40 to a final concentration of 0.5%. The NHS-ester chemical crosslinker BS³ (Bis(Sulpho-succinimidyl) Suberate, Sigma Ltd.) was prepared as a 30mM stock solution in 0.1M Hepes pH7.4 immediately prior to use and kept on ice throughout the procedure. Cross-linking reactions were carried out with final BS³ concentrations of; 0, 0.05, 0.2, 2.0, and 5.0mM and incubated at 4°C for 1 hour and then quenched with 50mM Tris-HCL pH8.5 for 5 minutes at 4°C. Proteins were analysed by SDS-6% PAGE under reducing and non-reducing conditions.

29/ PNGase F Treatment of Whole Cell Extracts.

293 cells were transfected with the appropriate T7-based vectors, as described above. Transfected cells were then harvested by washing and scrapping the cells into PBS, pelleting on a benchtop microcentrifuge and finally resuspended into x1 disruption buffer. The whole cell extract (WCE) was further denatured by boiling for 5 minutes, whereupon 1/10th volumes of reaction buffer (x10: 0.5M sodium phosphate pH7.5) and 10% Nonidet P40 (NP40) were added. 1000Us of PNGase F (reaction buffer, 10% NP40 and enzyme was purchased from New England Biolabs) were added to the reaction mixture, which was then incubated for 2hrs at 37°C. The extracts were analysed on SDS-PAGEs under reducing conditions followed by Western blotting using the anti-Pk monoclonal antibody and enhanced chemiluminescence detection (ECL) detection.

30/ Tunicamycin Treatment Of transfected cells.

Tunicamycin (Sigma) was prepared as a 1µgµl⁻¹ stock in DMSO and kept in a light-proof eppendorf at 4°C. 293 cells were transfected with the appropriate vectors as described

above except that tunicamycin at a final concentration of $2\mu\text{gml}^{-1}$ was added to the medium, 4hrs post-transfection. The cells were harvested 16hrs later and WCEs were analysed by a separation through SDS-PAGEs under reducing conditions, followed by Western blotting using the anti-Pk monoclonal antibody and ECL detection.

RESULTS

As the following results section contains a large amount of information on the construction and analysis of a number of plasmid vectors, a summary of the vectors used throughout, including some of their major features, is depicted in table 4.

SECTION 1 : ANALYSIS OF TRANSIENTLY EXPRESSED Pk-TAGGED SV5 GLYCOPROTEINS.

The vector initially chosen for transient expression of Pk-tagged glycoproteins in mammalian cells was the plasmid pGEM-3zf(+) (from now on referred to as pT7, figure 19). This vector possesses a promoter derived from the T7 bacteriophage, from which expression is driven by the T7 RNA-dependant RNA polymerase (T7-RDRP). The T7-RDRP was introduced into mammalian cell lines by infection with a recombinant vaccinia virus (vTF7-3) which encodes the T7-RDRP, as described by Elroy-Stein *et al* (1989). Expression from the vector was achieved by liposome-mediated transfection of the circular plasmid into the infected cells which were subsequently analysed for protein production 12 to 18 hrs post-transfection.

1.1) Construction of vectors for the expression of Pk-tagged HN & F glycoproteins.

1.1.1 / Construction of the vectors encoding an N or C-terminal Pk-tag.

Two oligonucleotide linkers, encoding the Pk- tag sequence (figures 16 & 17) for positioning at the N or C - terminus of an expressed protein, were cloned into pT7, producing the plasmids pT7ntPk and pT7ctPk respectively (figure 19). Both oligonucleotide linkers possessed restriction endonuclease (RE) recognition sites for the directional cloning of cDNAs and, where appropriate, translational initiation and termination signals. Plasmids

TABLE 4 Summary of plasmid vectors used throughout this thesis and some of their major features.

The symbol * refers to vectors constructed within this laboratory by either B. Precious or A. Bermingham. The symbol ** refers to vectors constructed outwith this laboratory. N/A - not applicable. The punctuation mark ? - unknown. The character (-> ->) refers to the favourable orientation, (-> <-) refers to the unfavourable orientation, as discussed in the text to follow. CMV refers to the human CMV major immediate-early enhancer/promoter element. Pol III refers to transcription initiating from an RNA pol III promoter. TRP refers to the tetracycline responsive promoter.

Vector	Promoter	Gene	EBNA-1/ EBV-ori	Orientation	SV40 - ori	Detectable gene expression
pT7ctPk	T7 ϕ 10	N/A	No	N/A	No	N/A
pT7ntPk	"	N/A	"	"	"	N/A
pT7HN	"	SV5 HN	"	"	"	Yes
pT7HNctPk	"	HNctPk	"	"	"	"
pT7HNntPk	"	HNntPk	"	"	"	"
pT7FctPk	"	FctPk	"	"	"	"
pT7 Δ HNctPk	"	Δ HNctPk	"	"	"	"
pT7 Δ HNntPk	"	Δ HNntPk	"	"	"	"
pT7 Δ FctPk	"	Δ FctPk	"	"	"	"
pCT7HNctPk	CMV & T7	HNctPk	No	N/A	Yes	Yes, T7 only
pCT7HNntPk	"	HNntPk	"	"	"	Yes, T7 only
pCT7LTA _g	"	LTA _g	"	"	"	Yes
pCT7 β gal	"	β gal.	"	"	"	Yes
pMR-tTa/neo *	CMV	tTa/neo	No	N/A	Yes	Yes
pEBV-tTa/neo *	"	tTa/neo	Yes	(-> ->)	No	"
pMR-tTa/hygB	"	tTa/hygB	No	N/A	Yes	"
pGEM-tTa/hygB	"	"	No	N/A	No	"
pEBV-tTa/hygB	"	"	Yes	(-> ->)	No	"
pTET/V *	TRP	SV5 V	No	N/A	No	Yes
pTETHNctPk	"	HNctPk	No	N/A	"	No
pTETHNntPk	"	HNntPk	No	N/A	"	Yes
pTEP4HNctPk	"	HNctPk	Yes	(-> ->)	"	No
pTEP4HNctPk	"	HNctPk	"	(-> <-)	"	No
pTEP4HNntPk	"	HNntPk	"	(-> ->)	"	Yes
pTEP4HNntPk	"	"	"	(-> <-)	"	"
pTEP8HNntPk	"	"	"	(-> ->)	Yes	"
pTEP8HNntPk	"	"	"	(-> <-)	Yes	"
pREP4HNctPk	RSV LTR	HNctPk	Yes	(-> <-)	No	No
pCEP4HNctPk	CMV	"	"	(-> <-)	"	"
pCEP4HNctPk	"	"	"	(-> ->)	"	"
pCEP8/P	"	SV5 P	"	(-> <-)	Yes	Yes
pCEP8/P	"	SV5 P	"	(-> ->)	Yes	Yes
pSVori **	SV40	LTA _g	No	N/A	Yes	Yes
pCMVluc **	CMV	Luciferase	"	"	?	Yes
pMHVA **	Pol III	VAI RNA	"	"	?	Yes


```

                                <-afl2->
      <--asc1-->                leu lys gly lys pro ile
CT GGG CGC GCC AAT AAT CTT AAG GGA AAG CCG ATC
GA CCC GCG CGG TTA TTA GAA TTC CCT TTC GGC TAG

pro asn pro leu leu gly leu asp ser thr OCH
CCA AAC CCT TTG CTG GGA TTG GAC TCC ACC TAA C
GGT TTG GGA AAC GAC CCT AAC CTG AGG TGG ATT GAGCT

```

Figure 16. DNA linker coding for the C-terminally positioned Pk-tag.

The schematic depicts a sequence for a double stranded oligonucleotide which encodes the Pk-tag oligopeptide derived from amino acids 95-108 of SV5 P&V proteins (**bold**). Amino acids in plain text are derived from the afl2 restriction site. After annealing, the oligonucleotide possesses a Pvu2 blunt - end and an Xho1 cohesive-end. Underlined region corresponds to the synthetic nonapeptide which was used in competition studies with the P and V proteins for binding to SV5-P-k antibody (Southern *et al*, 1991). Shown is the asc1 & afl2 sites used for the introduction of the gene of interest. The linker also possesses an ochre codon (OCH) as a signal for the termination of translation.

```

                                gly lys pro ile pro asn pro leu leu
CC GGG ACC ATG GGA AAG CCG ATC CCA AAC CCT TTG CTG
  C TGG TAC CCT TTC GGC TAG GGT TTG GGA AAC GAC

gly leu asp ser thr trp arg ala                leu lys OCH
GGA TTG GAC TCC ACC TGG CGC GCC TAG AAC CTT AAG TAA T
CCT AAC CTG AGG TGG ACC GCG CGG ATC TTG GAA TTC ATT AGATC
                                <--asc1-->        <-afl2->

```

Figure 17. DNA linker coding for the N-terminally positioned Pk-tag.

The schematic depicts a sequence for a double stranded oligonucleotide which encodes the Pk-tag oligopeptide derived from amino acids 95-108 of SV5 P&V proteins (**bold**). Amino acids in plain text are derived from the asc1 and afl2 restriction sites. After annealing, the oligonucleotide possesses an Xma1 cohesive-end and a Sal1 cohesive-end. Shown is the asc1 & afl2 sites used for the introduction of the gene of interest. Upstream of the ATG codon is the tri-nucleotide sequence ACC conferring an optimal Kozak sequence for translation initiation (Kozak, 1989). The linker also possesses an ochre codon (OCH) as a signal for the termination of translation.

pT7ntPk & pT7ctPk were sequenced to ensure the fidelity of the sequence (figure 18).

1.1.2 /Construction of the vectors encoding native SV5 HN and N or C - terminally Pk-tagged HN (figure 19).

The gene encoding the haemagglutinin/Neuraminidase protein from SV5 was amplified from the plasmid HN177 using PCR technology. Primers were used in which ascI & aflII RE recognition sequences were incorporated for the directional cloning of the PCR product into the pT7ctPk plasmid to produce the plasmid pT7HNctPk. The vector pT7HNntPk was constructed by the digestion of the plasmid pT7HNctPk with the REs ascI & aflII followed by a ligation of the excised HN gene into the similarly digested plasmid pT7ntPk. In addition, a vector termed pT7HN, encoding the native HN glycoprotein was constructed. pT7HN was produced in a manner similar to that of the pT7HNctPk plasmid, except that a TAA stop codon (which terminated translation after the last amino acid encoded by the HN gene) was incorporated into the PCR primer. Appropriate portions of the plasmids; pT7HN, pT7HNctPk & pT7HNntPk were sequenced to ensure fidelity (data not shown).

1.1.3 /Construction of the vector for the expression of C-terminally tagged SV5 F (figure 20).

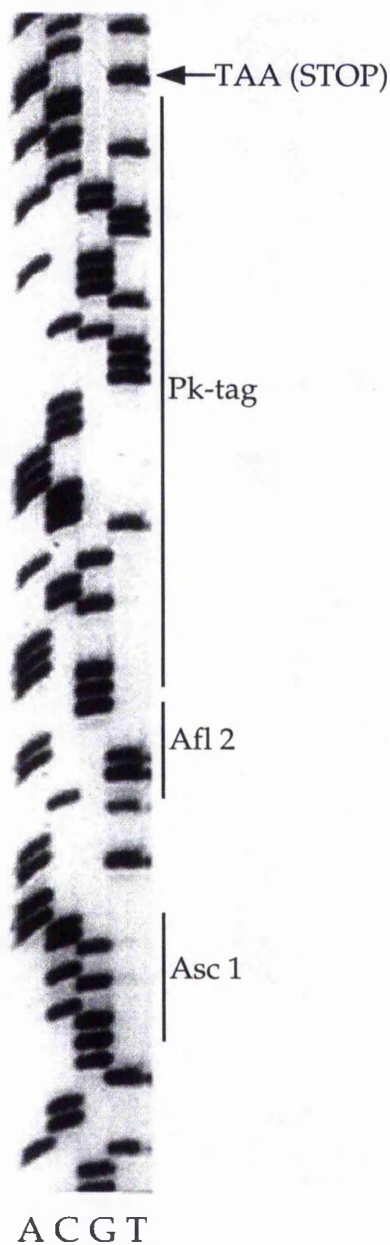
The gene encoding the fusion protein from SV5 was PCR-amplified from a plasmid containing the full length SV5 F gene (generously donated by R.A. Lamb) such that the PCR product contained a 5' ascI RE site and a 3' aflII RE site. The PCR product was subsequently digested with ascI & aflII and ligated into the similarly digested pT7HNctPk plasmid to produce the expression vector, pT7FctPk. The F gene was sequenced to ensure fidelity of the sequence (data not shown).

1.2 Transient protein production from the vectors which encode Pk-tagged HN & F glycoproteins.

1.2.1 /Expression of Pk-tagged SV5 HN & F.

Infection of 293 cells with vaccinia virus encoding the T7-RDRP followed by the subsequent transfection with the expression vectors was as described in the methods section.

A) CT Pk-tag



B) NT Pk-tag

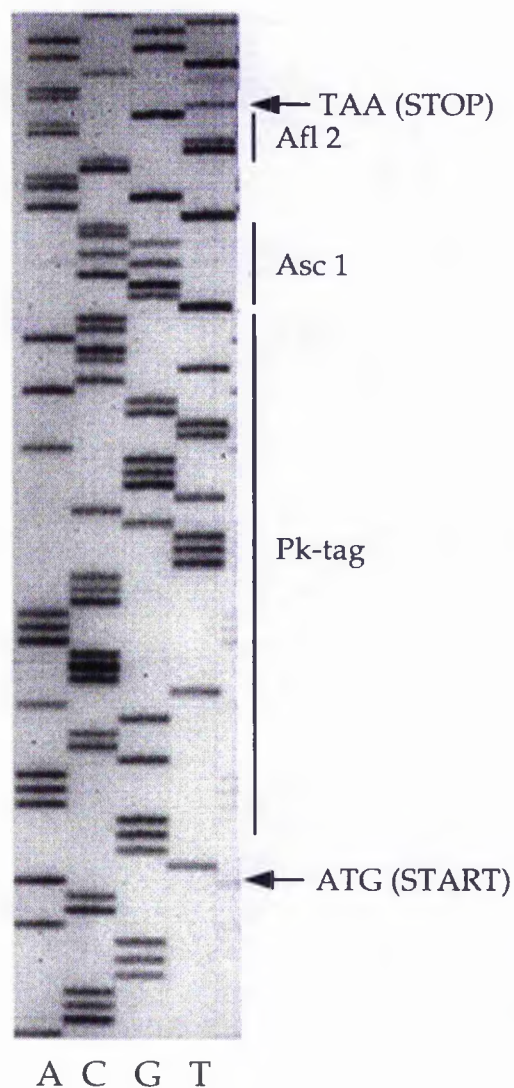


Figure 18: Sequence analysis of the CT & NT Pk-tag.

A schematic depicting the coding strand of the DNA sequence for the CT (A) & NT (B) Pk-tag oligopeptide epitope cloned into a pGEM plasmid to produce the vectors pT7ctPk & pT7ntPk. Also shown are the TAA translation termination (STOP) codons, ATG translation initiation (START) codon, and the recognition sequences for the restriction endonucleases Asc 1 (GGCGCGCC) and Afl 2 (CTTAAG). The sequence reads from the bottom to the top.

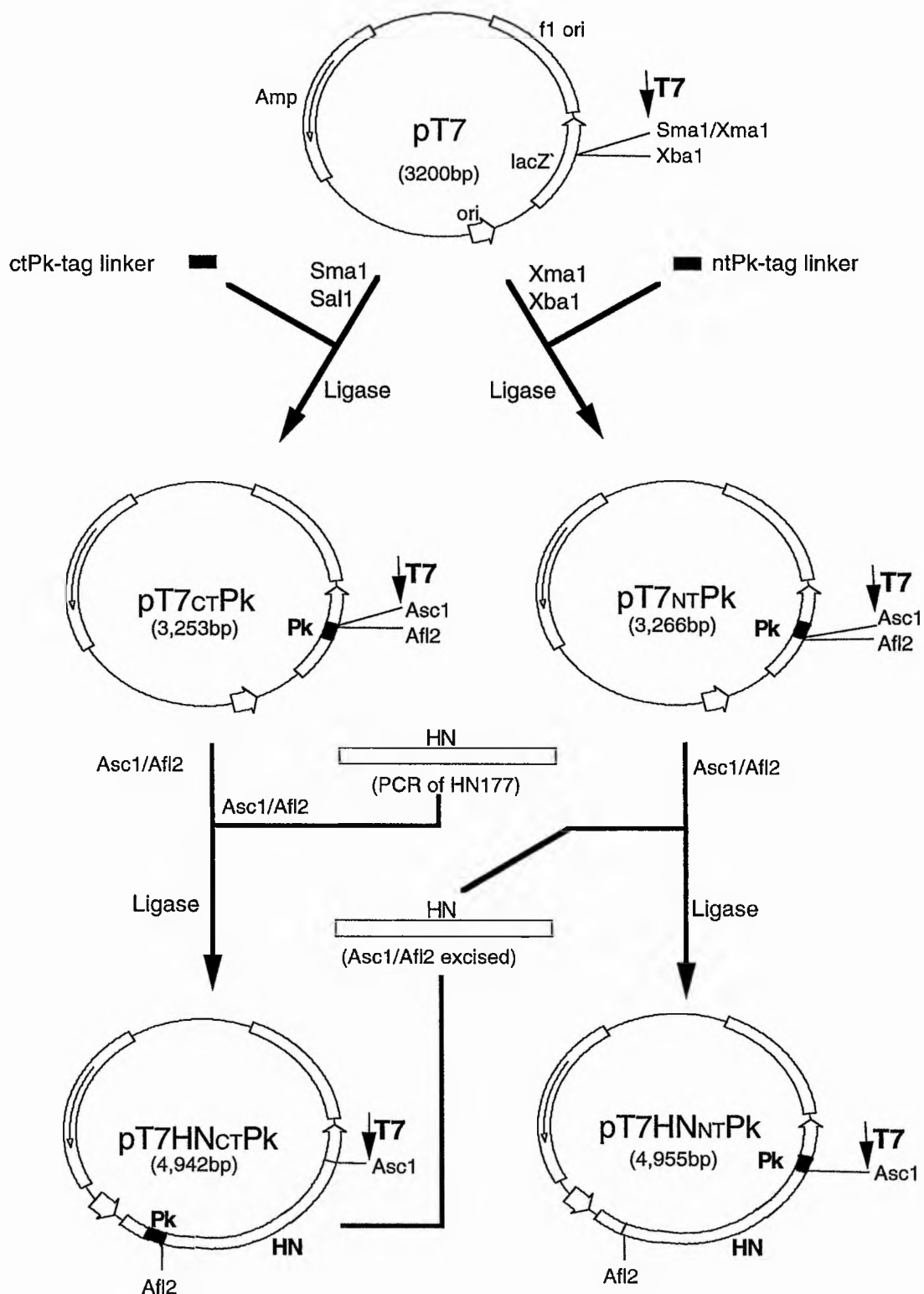


Figure 19: Construction of pT7HNntPk and pT7HNctPk vectors.

Two oligonucleotides encoding the Pk-tag sequence (figs. 16 & 17) were cloned into the vector pT7 producing the plasmids pT7ntPk or pT7ctPk, using the restriction endonuclease (RE) sites as shown. The gene coding for HN was amplified from the vector HN177 by PCR such that it could be directionally cloned into pT7ctPk, via the asc1 & afl2 RE sites, producing pT7HNctPk. To construct pT7HNntPk, the HN insert from pT7HNctPk was excised by asc1 & afl2 digestion and ligated into the similarly digested, pT7ntPk.

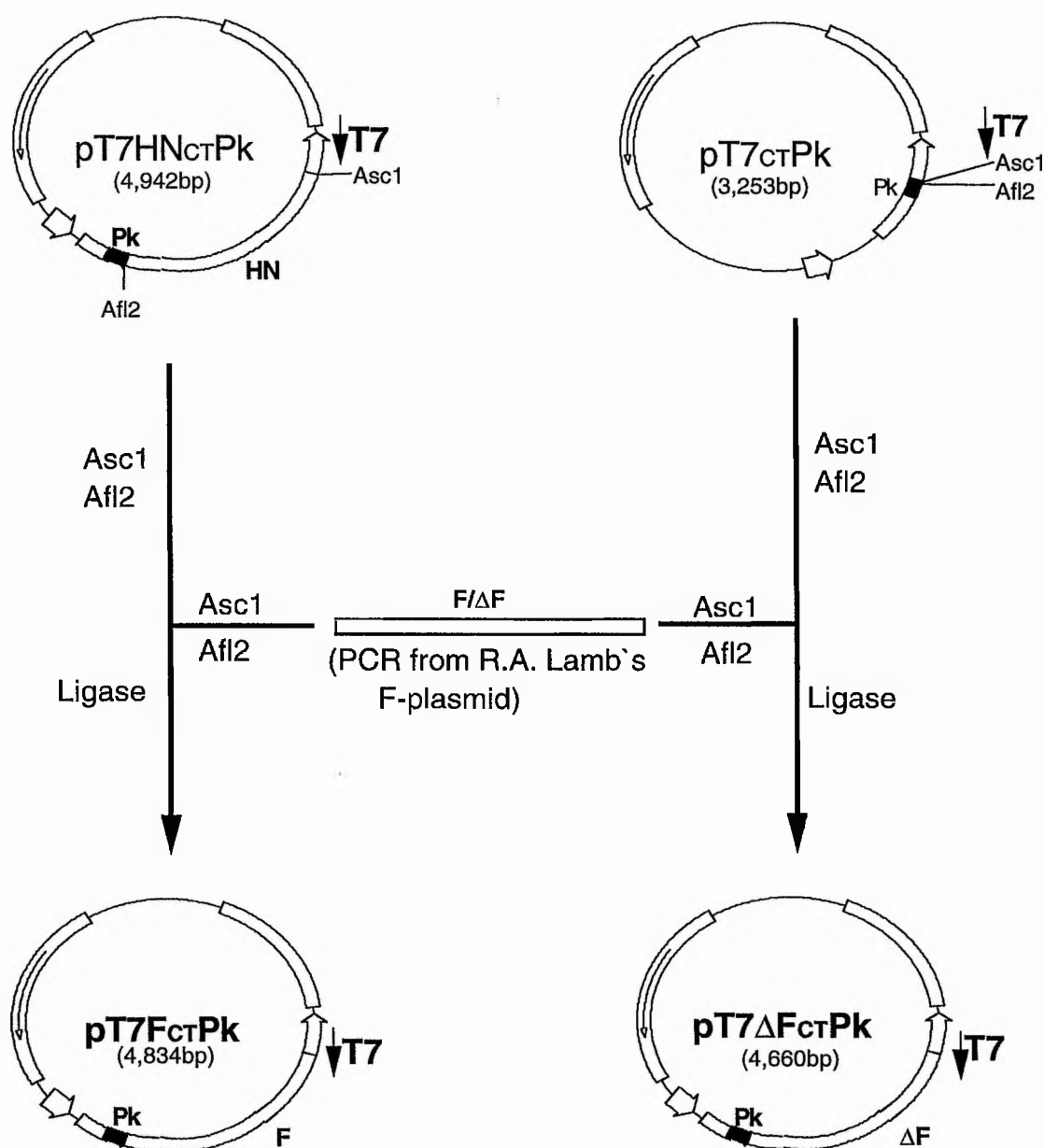


Figure 20: Construction of the vectors which encode the SV5 F glycoprotein and a truncated form, ΔF.

The gene encoding the SV5 F protein was amplified from a plasmid containing the entire SV5 F gene (donated by R.A. Lamb) by PCR such that it could be directionally cloned into pT7HNctPk, via the asc1 & afl2 RE sites, producing the plasmid pT7FctPk. To construct pT7ΔFctPk, the nucleotide sequence corresponding to amino acids 1 to 471 of the SV5 F gene was PCR amplified from pT7FctPk and the PCR product was digested with the REs Asc1 & Afl2 and ligated into the similarly digested vector pT7ctPk using the asc1 & afl2 RE sites, producing the vector pT7ΔFctPk. Transcription from both vectors was from the bacteriophage T7 ϕ 10 promoter.

18 hrs post-transfection, the cells transfected with the plasmids; pT7HNctPk, pT7HNntPk or pT7FctPk were washed once with PBS, detached from the plastic by scrapping, pelleted and then resuspended in x1 disruption buffer. After sonication, the whole cell extracts (WCEs) were separated through a 10% SDS-PAGE under reducing conditions, transferred to a PVDF membrane by electroblotting and probed with the anti-Pk mAb. Figure 21 shows that products of ~70Kda were detected from cells transfected with vectors encoding the N or C-terminally Pk-tagged HN (lanes 1 & 2), and products of ~ 66Kda & 56Kda were detected from cells transfected with the pT7FctPk construct (lane 3). No products were detected from mock transfections (lane 4).

1.2.2 *N-linked glycosylation analysis of Pk-tagged SV5 HN & F.*

N-linked glycosylation adds approximately 6-10 Kda to the molecular weight of the polypeptide backbone of native F and HN glycoproteins, O-linked glycosylation does not appear to be involved. To determine whether a protein has undergone the addition of N-linked sugar residues, the use of the enzyme PNGase F (peptide-N⁴-(N-acetyl- β -glucosaminyl)asparagine amidase), which hydrolyses the asparaginyl-oligosaccharide bond on the polypeptide, can be employed. From figure 21, indirect evidence for glycosylation of Pk-tagged HN and F could be inferred from the relative molecular weights of the detected proteins.

To directly ascertain whether Pk-tagged HN and F were N-linked glycosylated, WCEs were treated with PNGase F as described in the methods section. Treated and untreated WCEs were separated through a 10% SDS-PAGE under reducing conditions, transferred to a PVDF membrane by electroblotting and probed with the anti-Pk mAb. Figure 22 shows that when WCEs from pT7HNctPk, pT7HNntPk & pT7FctPk transfected cells were treated with the enzyme PNGase F, the molecular weights of the expressed products were decreased by approximately 6-10 Kda directly indicating that the Pk-tagged glycoproteins had undergone N-linked glycosylation.

An alternative method of determining whether N-linked glycosylation of a glycoprotein occurs is by employing reagents which block the formation of the dolichol

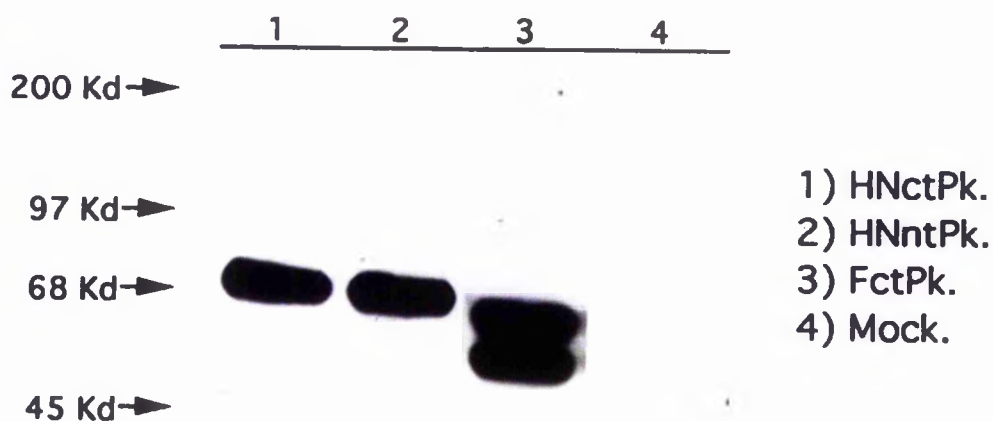


Figure 21: Transient expression of Pk-tagged HN and F glycoproteins and their antigenicity to the anti-Pk mAb.

vTF7-3 infected 293 cells were transfected with the plasmids; pT7HNctPk (lane 1), pT7HNntPk (lane 2), pT7FctPk (lane 3) or were mock transfected (lane 4). 16hrs post-transfection WCEs were prepared in disruption buffer and aliquots separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb. Detection was by ECL.



Figure 22: PNGase F analysis of Pk-tagged HN, F and ΔF.

vTF7-3 infected 293 cells were transfected with the vectors; pT7HNctPk (lanes 1 & 2), pT7HNntPk (lanes 3 & 4), pT7FctPk (lanes 5 & 6) and pT7ΔFctPk (lanes 7 & 8). 16hrs post-transfection, WCEs were prepared in disruption buffer. Half of the samples were treated with the enzyme PNGase F, as detailed in the methods. Untreated/treated extracts were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb. Detection was by ECL.

phosphate-linked oligosaccharide donor, thereby preventing N-linked glycosylation of the polypeptide backbone. An example of such an agent is the antibiotic tunicamycin. Treatment of transfected cells with tunicamycin was carried out as detailed in the methods section. Transfected cells showed a similar decrease of approximately 6-10 Kda in the relative molecular weights of the expressed proteins when tunicamycin was present, (data not shown).

1.2.3 /Localisation of Pk-tagged SV5 HN & F.

Having shown that Pk-tagged forms of HN and F are N-linked glycosylated, as previously shown for native HN and F, the cellular localisation of expressed Pk-tagged HN and F was determined. Native SV5 HN and F are located in the plasma-membrane on the surface of infected or appropriately transfected cells. Therefore, it was investigated using indirect immunofluorescence, whether or not the Pk-tagged forms of the membrane-bound HN and F are also localised to the surface of transfected cells. Furthermore, the correct topology of Pk-tagged HN & F within the plasma-membrane, was investigated by analysis of cells which had either been fixed only, or fixed then permeabilised.

vTF7-3 infected 293 cells, grown on glass coverslips in 35mm petri dishes, were transfected with plasmids expressing; FctPk, HNctPk, HNntPk, native HN and, as a control, β -galactosidase, and processed for indirect immunofluorescence as detailed in the methods section. Figure 23 shows that native HN was detected on the surface of transfected cells by a pool of anti-HN mAbs, but not with the anti-Pk mAb, as expected. N-terminally Pk-tagged HN was detected on the surface of the majority of transfected cells using the pool of anti-HN mAbs (figure 24), with only a small number of cells showing a specific staining with the anti-Pk mAb [the most likely reason that some cells did show specific staining with the anti-Pk mAb, is due to a few cells which have lost their plasma-membrane integrity. This can be observed with the staining for the cytoplasmic protein β -galactosidase (figure 25) where transfected monolayers, which were not permeabilised, possessed a small number of cells with specific intracellular staining]. Permeabilisation of the pT7HNntPk transfected cells did result in the detection of N-terminally Pk-tagged HN in the majority of transfected cells using

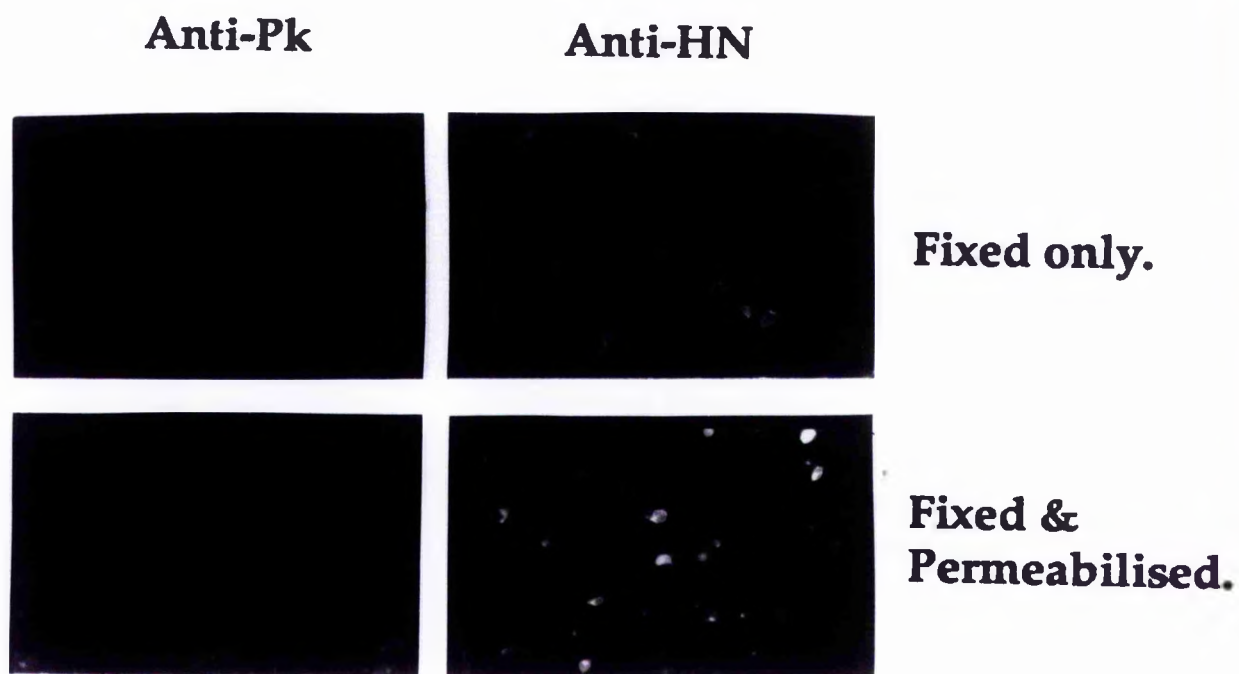


Figure 23: Immunofluorescence analysis of untagged HN glycoprotein.

vTF7-3 infected 293 cells were transfected with the vector pT7HN. 12hrs post-transfection, cells were fixed only or fixed & permeabilised and incubated with anti-Pk or a pool of anti-HN mAbs (1b, 4b & 5d) followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant.

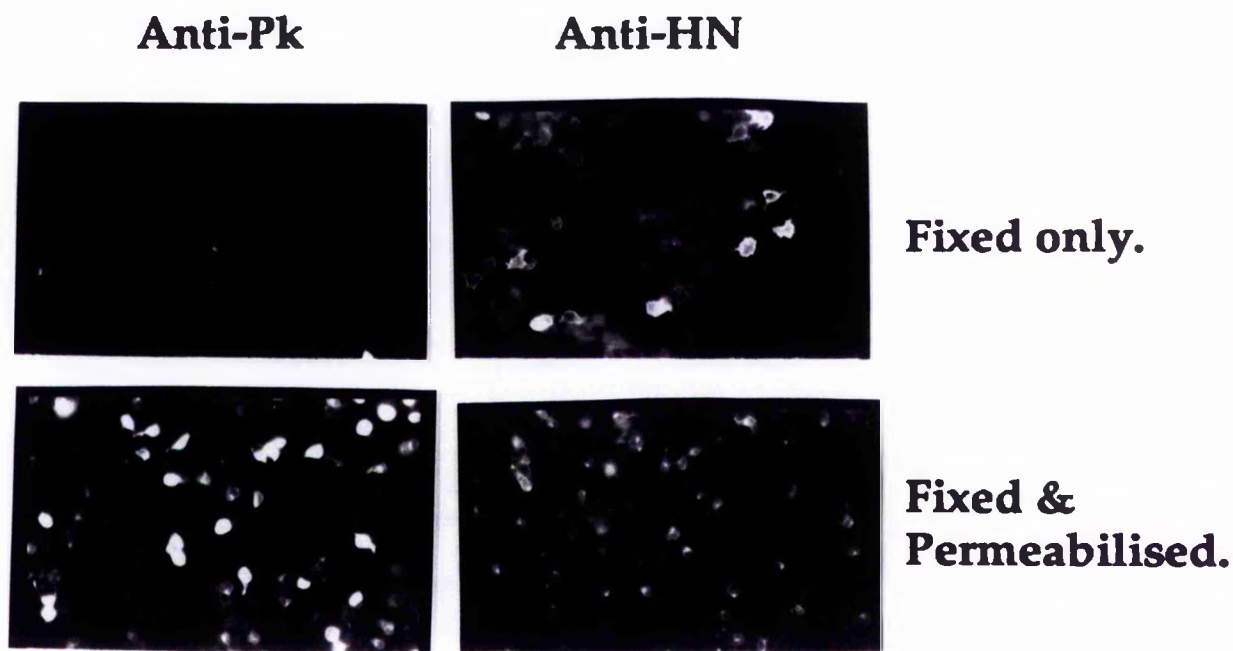


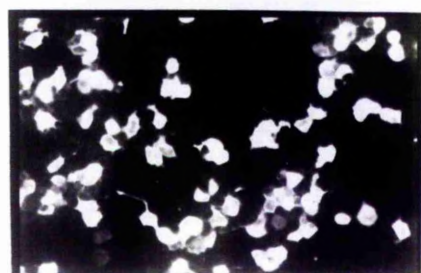
Figure 24: Immunofluoresence analysis of N-terminally Pk-tagged HN glycoprotein.

vTF7-3 infected 293 cells were transfected with the vector pT7HNntPk. 12hrs post-transfection, cells were fixed only or fixed & permeabilised and incubated with anti-Pk or a pool of anti-HN mAbs (1b, 4b & 5d) followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant.

Anti- β gal



Fixed only.



**Fixed &
Permeabilised.**

Figure 25: Immunofluoresence analysis of 293 cells expressing β -galactosidase.

293 cells infected with vaccinia virus vTF7-3 were transfected with the vector pCT7 β gal. 16 hrs post-transfection, cells were either fixed-only or fixed & permeabilised then incubated with an anti- β gal mAb followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant.

the anti-Pk mAb (figure 24), thus confirming that the Pk epitope was concealed within the cytoplasm, as expected.

Cells transfected with the vector expressing the C-terminally Pk-tagged HN were specifically stained for surface immunofluorescence using both the pool of anti-HN mAbs and the anti-Pk mAb, thus confirming the presence of the Pk-tag epitope on the exterior of the cell, as expected (figure 26) Permeabilisation of the above transfected cells, also resulted in the specific intracellular staining with both the pool of anti-HN mAbs and the anti-Pk mAb.

The majority of pT7FctPk transfected cells, which were fixed only, could be specifically stained with the anti-F mAb, with only a small number showed specific staining with the anti-Pk mAb, figure 27. The reason for this detection is most likely due to a similar situation as described above for the ntPk-tagged HN construct. Permeabilisation of the transfected cells resulted in the specific staining of the majority of the transfected cells, by both the anti-F & anti-Pk mAbs, thus confirming the presence of the Pk epitope in the interior of the cell, as expected (figure 27).

1.2.4 /Oligomeric structure of Pk-tagged SV5 HN & F.

As previously mentioned, the HN glycoprotein of SV5 forms a mixture of covalently and non-covalently linked homotetramers composed of disulphide-linked dimers. The SV5 F glycoprotein appears to form non-covalently linked homotrimers. Therefore, it was investigated to determine whether or not the Pk-tagged forms of these glycoproteins also formed similar oligomeric structures. This was achieved by using non-reducing SDS-PAGE and chemical cross-linking analysis for the study of the oligomeric structures of Pk-tagged HN and F, respectively. Figure 28 shows WCEs from pT7HNntPk and pT7HNctPk transfected cells which have been separated through an SDS-PAGE under reducing (β -mercaptoethanol present in the disruption buffer) or non-reducing (β -mercaptoethanol absent) conditions, electroblotted and probed with the anti-Pk mAb. A single band representing monomeric Pk-tagged HN was detected in the reduced sample, and to a lesser extent in the non-reduced sample. However, diffuse bands corresponding to molecular weights of approximately 140 and 280 Kda were detected in the non-reduced sample, but

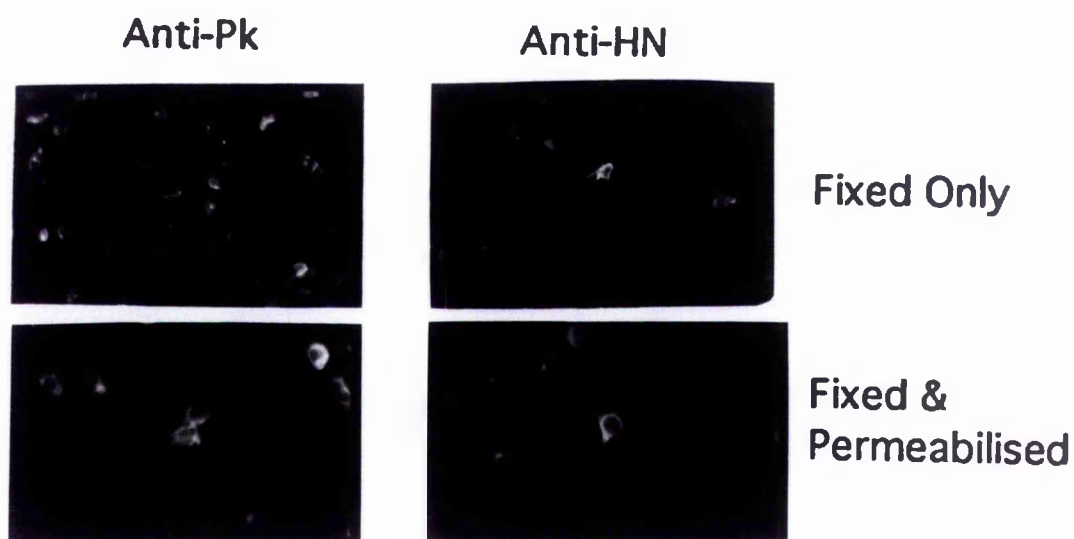


Figure 26: Immunofluoresence analysis of C-terminally Pk-tagged HN glycoprotein.

vTF7-3 infected 293 cells were transfected with the vector pT7HNctPk. 12hrs post-transfection, cells were fixed only or fixed & permeabilised and incubated with anti-Pk or a pool of anti-HN mAbs (1b, 4b & 5d) followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant.

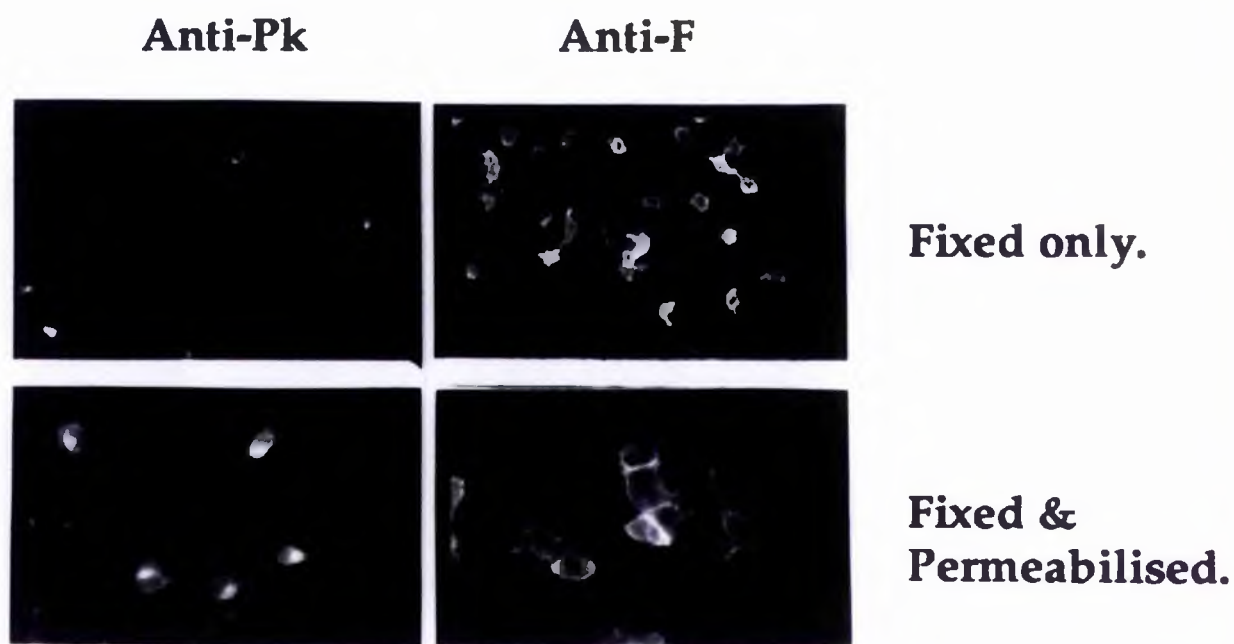


Figure 27: Immunofluoresence analysis of C-terminally Pk-tagged F glycoprotein.

vTF7-3 infected 293 cells were transfected with the vector pT7FctPk. 12hrs post-transfection, cells were fixed only or fixed & permeabilised and incubated with anti-Pk or anti-F (mAbs followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant.



Figure 28: Analysis of the oligomeric structure of Pk-tagged SV5 HN.

vTF7-3 infected 293 cells were transfected with the vectors pT7HNctPk (lanes 1 & 4), pT7HNntPk (lanes 2 & 5) or were mock transfected (lanes 3 & 6). 16hrs post-transfection WCEs were prepared in disruption buffer in the presence or absence of the reducing agent β -mercaptoethanol (B.M.E.). The WCEs were separated through a 10% SDS-PAGE, electroblotted and probed with the anti-Pk mAb. Detection was by ECL.

were absent from the reduced sample. These bands correspond to dimeric and tetrameric forms of the C and N-terminally Pk-tagged HN.

Chemical cross-linking reactions performed on WCEs from pT7FctPk transfected cells were carried out as detailed in the methods section. Results from the chemical cross-linking analysis of C-terminally Pk-tagged F are shown in figure 29 and demonstrated that, under non-reducing conditions (panel A), specific bands corresponding to monomers (~66Kda), dimers (~132Kda) and trimers (~198Kda) were observed. Increasing the concentration of the BS³ cross-linking reagent from 0-5mM influenced the molecular weight of the bands detected, ranging from predominantly monomers to predominantly trimers. Analysis of the cross-linked extracts under reducing conditions (panel B) demonstrated the reduction of the F_O subunits to F₁ & F₂ subunits.

1.2.5 /Antigenicity of Pk-tagged SV5 HN & F to mAbs raised against native SV5 HN & F, which recognise conformationally-dependent epitopes.

Preliminary characterisation of the Pk-tagged glycoproteins has suggested that the presence of the Pk-tag did not significantly affect the processing of the Pk-tagged glycoproteins through the mammalian-cell exocytotic pathway. However, with a view to producing Pk-tagged glycoproteins for immunisation purposes, it was imperative to determine whether the Pk-tagged glycoproteins bind to conformationally-dependant mAbs, which were raised against the native glycoproteins. This was investigated by performing immunoprecipitation experiments.

A panel of 18 mAbs to the SV5 HN glycoprotein and one mAb to the SV5 F glycoprotein have been produced and their specificity determined in our laboratory (Randall *et al.*, 1987). vTF7-3 infected 293 cells were transfected with the plasmids pT7HNntPk and pT7HNctPk. 16hrs post-transfection, soluble cell extracts (SCEs) were prepared and immunoprecipitations were performed as detailed in the methods section, using the anti-Pk mAb (positive control), anti-DNP mAb (negative control) and 8 anti-HN mAbs; 1b, 5d, 5a, 4a, z, 4b, 4d & 4f, (where a number/letter refers to the antigenic group that a particular mAb was assigned to, see Randall *et al.*, 1987).

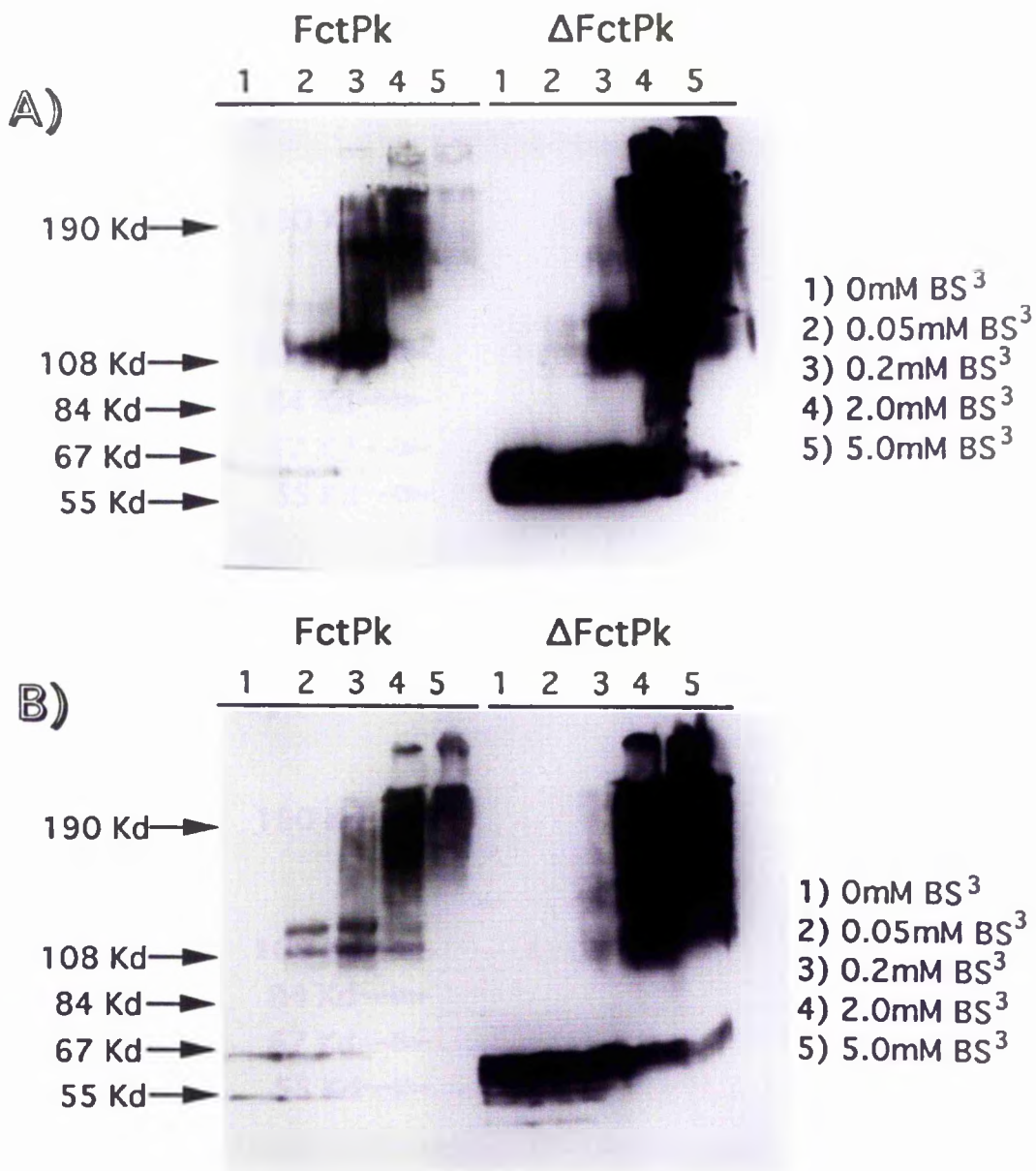


Figure 29: Oligomeric structure of the Pk-tagged form of SV5 F and of a truncated form Δ F, by chemical cross-linking analysis.

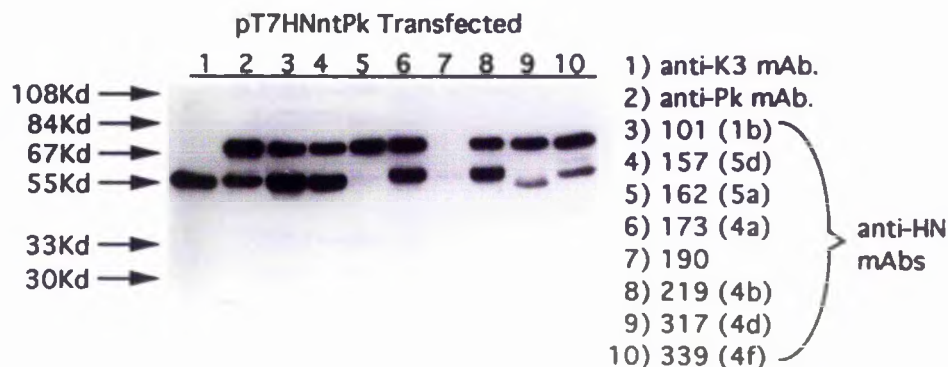
vTF7-3 infected 293 cells were transfected with the plasmids pT7FctPk and pT7 Δ FctPk. 12 hrs post-transfection, SCEs were prepared by lysis of the cells with 1% NP40 in PBS, followed by a 1 minute centrifugation at 10,000g. The supernatant was divided into 5 aliquots one of which was left untreated (lane 1). The other 4 aliquots were incubated, for 1 hr on ice, with increasing concentrations of the NHS-ester cross-linking reagent BS3 (lanes 2 to 5). The cross-linking reaction was quenched by the addition of Tris-CL pH 8.0 to 50mM. The SCEs were then made upto x1 disruption buffer (minus β -mercaptoethanol) and separated through a 6% SDS-PAGE under non-reducing conditions (panel A) or reducing conditions (plus β -mercaptoethanol, panel B), electroblotted and probed with the anti-Pk mAb. Detection was by ECL.

Figure 30 shows that the N and C-terminal Pk-tagged forms of the HN glycoprotein (corresponding to bands detected at ~70Kda) were recognised by 7 out of the 8 anti-HN mAbs tested. Lane 7 in panel A and lane 10 in panel B were negative because that particular anti-HN mAb (190) was not recognised by protein A present on the fixed and killed *S. aureus*. Bands present at ~55Kda correspond to the heavy chains of the mAbs employed in the immune precipitation. An immunoprecipitation analysis of the reactivity of the anti-F mAb with C-terminally Pk-tagged F was not performed because, from immunofluorescence data, figure 27, it was clear that the anti-F mAb did recognise the Pk-tagged form of F. Quantification of the reactivity of the seven anti-HN mAbs to the N and C-terminally Pk-tagged forms of the HN glycoprotein, was investigated using phosphoimaging technology. This was achieved by separating the immunoprecipitated samples through a 10% SDS-PAGE, electroblotting and probing with the anti-Pk mAb as the primary detection agent, with ¹²⁵Iodine-labelled anti-mouse immunoglobulin being used as the secondary detection reagent. The results were then analysed using a Fuji BAS 1000 phosphoimager. However, consistent data was not produced by this approach (data not shown).

1.3 / Construction of vectors for the expression of truncated Pk-tagged HN & F.

From the preceding results, the addition of the 14 amino acid oligopeptide to the termini of recombinantly produced SV5 HN & F glycoproteins did not significantly affect their processing through the mammalian-cell exocytotic pathway. Hence, N-linked glycosylation, oligomerisation, surface localisation and antigenicity to monoclonal antibodies raised against the native glycoproteins, were all retained. For the purpose of large-scale recombinant protein production, it may be beneficial to produce recombinant membrane-bound glycoproteins which are in a secretable form for increased productivity and ease of purification. Therefore, the genes encoding the Pk-tagged HN & F glycoproteins were altered such that the expressed proteins should be capable of being secreted from mammalian cells.

A)



B)

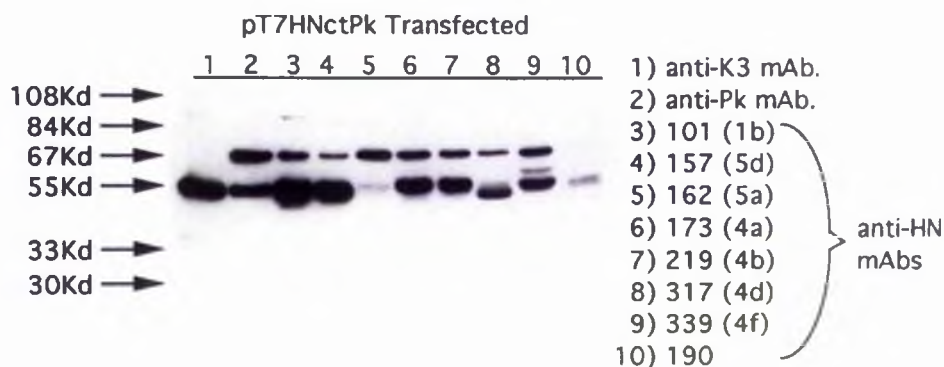


Figure 30: Immunoprecipitation analysis demonstrating the antigenicity of transiently expressed Pk-tagged HN glycoproteins for a panel of conformationally sensitive anti-HN mAbs.

vTF7-3 infected 293 cells were transfected with the plasmids pT7HNntPk (A) or pT7HNctPk (B). 16hrs post-transfection SCEs were prepared in immunoprecipitation buffer and aliquots were immunoprecipitated individually with 8 different anti-HN mAbs (lanes 3-10), anti-Pk mAb (lane 2) or anti-K3 mAb (-ve control, lane 1). Immunoprecipitates were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb. Detection was by ECL. (Note - mAb 190, panel A, is not bound by protein A and therefore HN was not pulled down. Bands at ~55Kda correspond to immunoglobulin heavy chain).

1.3.1 /Construction of vectors for the expression of truncated Pk-tagged HN (figure 32).

In collaboration with Dr. Mike Mackett, (Paterson Institute for Cancer Research, Univ. of Manchester) the plasmids pT7ctPk and pT7ntPk were altered by the addition of oligonucleotides encoding amino acids 1-22 (figure 31) of the EBV major surface glycoprotein gp360/220, the sequence of which encodes a putative type I cleavable signal sequence. A truncated HN gene from base 109 (amino acid 37, (see appendix A for the ORF sequence of SV5 HN)) of the ORF up to and including the C-terminal Pk-tag, was PCR-amplified from the plasmid pT7HNctPk. The PCR product was digested with the REs asc1 & Hind3 and ligated into the similarly digested vector pT7ctPk/ss (figure 32, **B**) to produce pT7ΔHNctPk (figure 32, **C**).

ATG GAG GCA GCC TTG CTT GTG TGT CAG TAC ACC
met gl u ala ala leu leu val cys gl n tyr thr

ATC CAG AGC CTG ATC CAT CTC ACG GGG GTA CCC
ile gl n ser leu ile his leu thr gly val pro

Figure 31. Putative cleavable signal sequence from the EBV type I glycoprotein gp360/220

To construct the vector pT7ΔHNntPk, the sequence comprising the HN ORF and N-terminal Pk-tag was excised from the plasmid pT7HNntPk using the REs Kpn1 & Hind3 and ligated into the similarly digested vector pT7ntPk/ss to produce pT7HNntPk/ss. A truncated HN gene, from base 109 (amino acid 37) to the 3' end of the HN ORF, was PCR amplified from plasmid pT7HNctPk. The PCR product was digested with asc1 & afl2 REs and ligated into the similarly digested plasmid pT7HNntPk/ss to produce pT7ΔHNntPk (figure 32, **C**). The plasmids pT7ΔHNntPk and pT7ΔHNctPk were sequenced at the 5' and 3' ends of the inserted genes to ensure the fidelity of the sequence (data not shown).

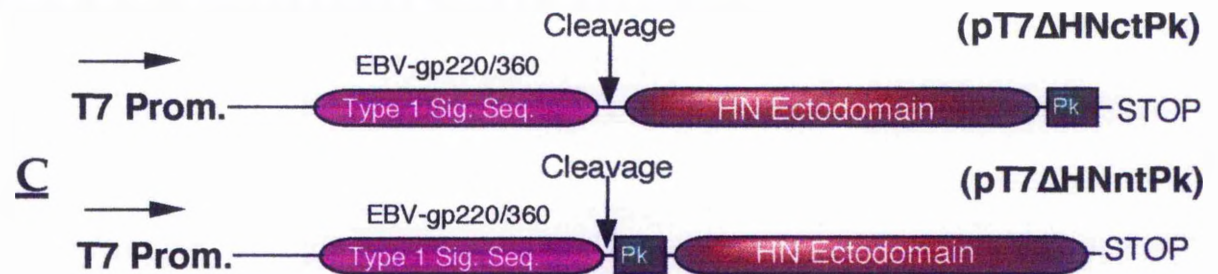
1.3.2 /Construction of vectors for the expression of truncated Pk-tagged F (figure 20).

To construct pT7ΔFctPk, a truncated version of the F gene lacking the nucleotides coding for the cytoplasmic tail and C-terminal membrane-spanning-domain (bases 1414 to

Schematic of Secretable forms of Type 1&2 Glycoproteins.



A) Secretable (Type2) Glycoprotein



B) Secretable (Type1) Glycoprotein

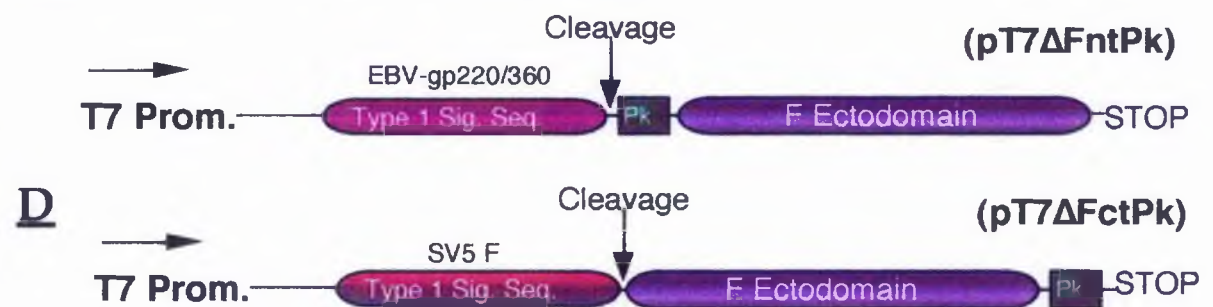


Figure 32. A Schematic depicting potentially secretable forms of type I and type II glycoproteins respresented by the SV5 F and HN glycoproteins respectively.

Plasmids pT7ntPk & pT7ctPk (**A**) were modified by the addition of a putative cleavable signal sequence from the Epstein-Barr virus surface glycoprotein gp 220/360, a type I glycoprotein (**B**). A truncated version of the SV5 HN glycoprotein, lacking its N-terminal signal sequence and membrane-spanning-domain, was cloned into the modified plasmids to produce vectors which should putatively secrete Pk-tagged truncated HN (**C**). A truncated version of the SV5 F glycoprotein, lacking its C-terminal cytoplasmic tail and membrane-spanning-domain, was cloned into the plasmid pT7FctPk to produce a vector which should putatively secrete C-terminally Pk-tagged F (**D**). The plasmid for the expression of N-terminally Pk-tagged truncated F was not constructed.

1587, amino acids 472 to 529 (see appendix B for the ORF sequence of SV5 F) was PCR amplified from the plasmid pT7FctPk. The PCR product was digested with the REs ascI & aflII and ligated into the similarly digested plasmid pT7FctPk to produce the vector pT7ΔFctPk. The full length truncated F gene was sequenced to ensure fidelity of the sequence (data not shown).

1.4 Transient protein production from the vectors which encode truncated Pk-tagged SV5 HN & F.

1.4.1 /Expression and localisation of the truncated N & C-terminally Pk-tagged SV5 HN.

1x10⁶ 293 cells were infected with the vaccinia virus recombinant vTF7-3 and subsequently transfected with 3μg of the expression vectors; pT7HNctPk, pT7HNntPk, pT7ΔHNntPk & pT7ΔHNctPk, as detailed in the materials and methods section. 16 hrs post-transfection WCEs were prepared in x1 disruption buffer and separated through a 10% SDS-PAGE under reducing conditions, transferred to a nitrocellulose membrane by electroblotting and probed with the anti-Pk mAb. Figure 33 shows that products of ~70Kda were detected from cells transfected with vectors encoding the N or C-terminally Pk-tagged HN and that doublet bands of ~55-60Kda were detected from WCEs transfected with the truncated vectors, with a minor band occurring at ~68Kda from transfected pT7ΔHNntPk but not pT7ΔHNctPk. A band of < 30Kda was detected from WCEs of cells transfected with the truncated C-terminally tagged HN, which is also present from the extract of cells transfected with the full length C-terminally tagged HN.

To ascertain whether any ΔHN was secreted into the external milieu, vTF7-3 infected 293 cells were transfected with the vectors pT7ΔHNntPk & pT7ΔHNctPk. 18hrs post-transfection the medium was aspirated from the cells, centrifuged at 13Krpm for 1 minute in a benchtop microcentaur and an immunoprecipitation carried out using the anti-Pk mAb together with a pool of anti-HN mAbs (1b, 4b & 5d), as detailed in the methods section. Furthermore, the cellular fraction was washed twice with PBS, detached from the plastic by scraping, pelleted, resuspended in immunoprecipitation buffer and incubated with a pool of anti-HN mAbs (1b, 4b & 5d). This was performed to ascertain whether the minor band at ~68Kda reacted with conformationally sensitive anti-HN mAbs. From figure 34, no specific

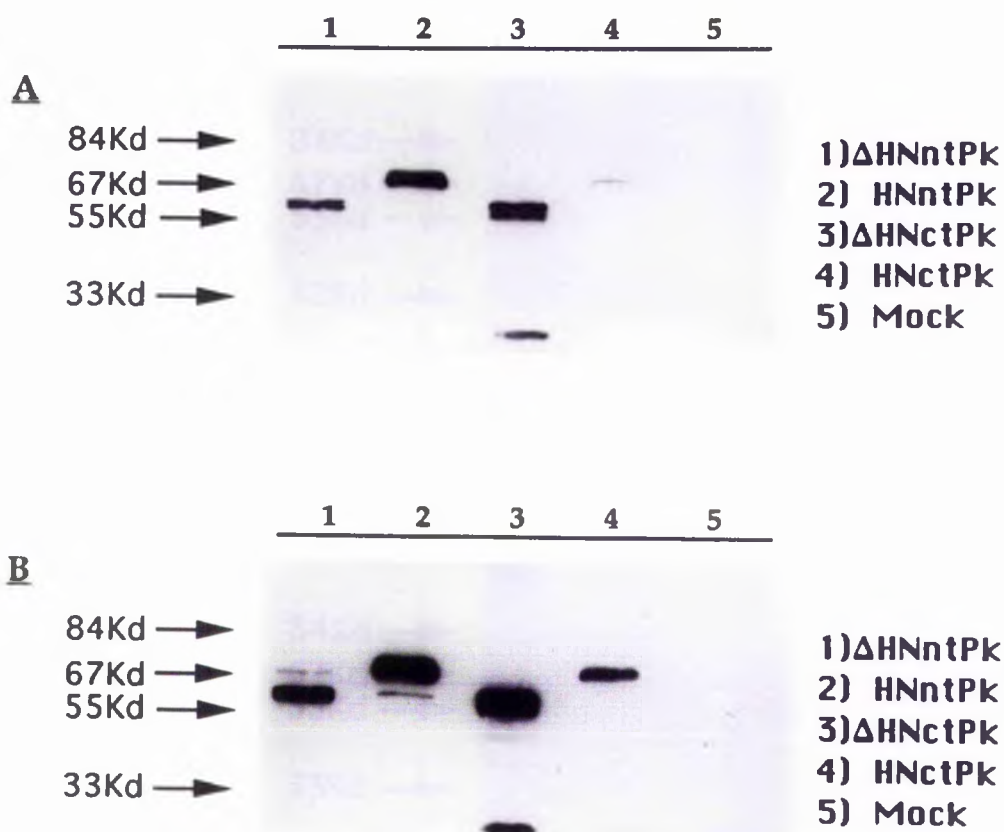


Figure 33: Transient expression of Pk-tagged truncated HN (Δ HN) glycoproteins and their antigenicity to the anti-Pk mAb.

vTF7-3 infected 293 cells were transfected with the plasmids; pT7 Δ HNntPk (lane 1), pT7HNntPk (lane 2), pT7 Δ HNctPk (lane 3), pT7HNctPk (lane 4) or were mock transfected (lane 5). 16hrs post-transfection WCEs were prepared in x1 disruption buffer and aliquots separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by detection using ECL. Panel A corresponds to a short exposure (10sec.). Panel B corresponds to a longer exposure (2 mins.).

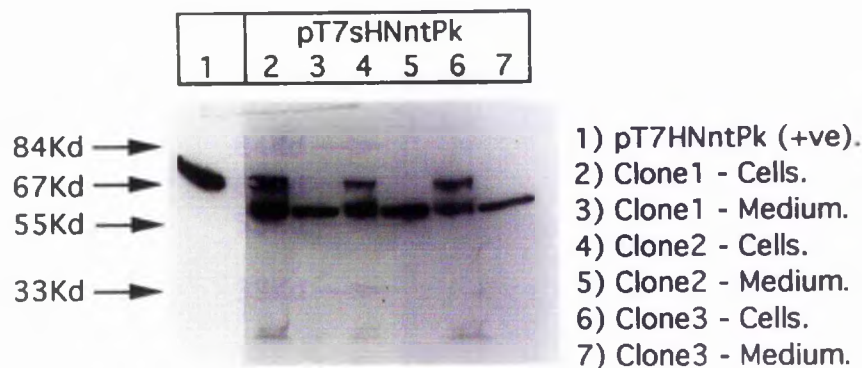


Figure 34: Immunoprecipitation (IP) analysis of ntPk-tagged truncated HN (Δ HN) glyco- protein.

vTF7-3 infected 293 cells were transfected with three different clones of the vector pT7 Δ HNntPk. 16hrs post-transfection the medium was aspirated, centrifuged to remove cellular debris and IP with anti-Pk mAb and a pool of anti-HN mAbs (1b, 4b & 5d), (lanes 3, 5 & 7). The cells were washed twice with PBS and SCE were prepared in immunoprecipitation buffer and immunoprecipitated as above (lanes 2, 4 & 6). The precipitated samples were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb. Detection was by ECL. A WCE from cells transfected with pT7HNntPk was also ran for a positive control. Note - exposure time for the immunoprecipitates was 15 mins. while only 20 secs. for the positive control (bands at ~55Kda correspond to immunoglobulin heavy chain).

bands were observed from the immunoprecipitation analysis performed on the supernatant fractions but a band of ~68Kda (bands at ~55Kda correspond to immunoglobulin heavy chain) was detected from the cellular fraction of pT7ΔHNntPk transfected cells, but not the pT7ΔHNctPk transfected cells (data not shown for pT7ΔHNctPk).

1.4.2 /Expression and localisation of the truncated C-terminally Pk-tagged SV5 F.

1x10⁶ 293 cells were infected with vTF7-3 and then transfected with 3μgs of the expression vectors pT7FctPk & pT7ΔFctPk, as detailed in the methods section. 16 hrs post-transfection the supernatant fractions were aspirated from the cells, centrifuged at 13Krpm for 1 minute in a benchtop microcentaur and the medium was concentrated eight-fold using an amicon 30Kda cut-off filter. The cellular fraction was washed twice with PBS, detached from the plastic by scrapping, pelleted and then resuspended in x1 disruption buffer. After sonication, the whole cell extracts (WCEs) along with the unconcentrated and concentrated supernatant samples, were separated through a 10% SDS-PAGE under reducing conditions, transferred to a PVDF membrane by electroblotting and probed with the anti-Pk mAb. Figure 35 shows that doublet bands of ~60-63Kda were visible from the cellular fraction after a short exposure, and a further 2 bands of ~58 & 53Kda were detected after a longer exposure. No bands were observed from the supernatant samples.

1.4.3 / N-linked glycosylation analysis of truncated Pk-tagged SV5 HN & F.

From figure 33, it could be inferred from the molecular weight of the majority of the expressed molecules, that the truncated Pk-tagged HN proteins have not undergone extensive post-translational modification. However, indirect evidence for the modification of a small proportion of the truncated ntPk-tagged HN, could be inferred from the relative molecular weight of a minor protein species detected at ~ 68 Kda (figure 33, lane1). This was not observed for the truncated Pk-tagged F as shown in figure 35, where the majority of the expressed molecules were of a molecular weight which would indicate some form of post-translational modification has occurred.

To directly ascertain whether the truncated forms of Pk-tagged HN & F were N-linked glycosylated, WCEs were treated with PNGase F as described in the methods section.

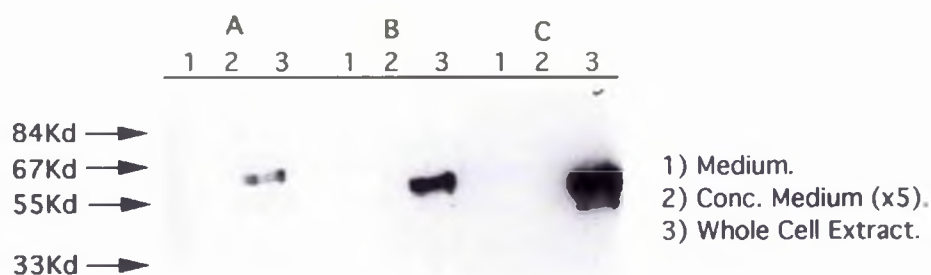


Figure 35: Transient expression and localisation of Pk-tagged truncated F (ΔF) glycoprotein .

vTF7-3 infected 293 cells were transfected with the plasmid pT7 ΔF ctPk. 16hrs post-transfection the medium was aspirated and centrifuged to remove cellular debris. An aliquot of this unconcentrated sample was made up to x1 disruption buffer. The remainder of the medium was concentrated 8-fold using a Amicon 30Kda cut-off membrane, where another aliquot was made up to x1 disruption buffer. WCEs were also prepared in x1 disruption buffer (after washing the cells twice with PBS). Samples from each preparation were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by detection using ECL. Panel A corresponds to a short exposure (20sec.). Panel B corresponds to a longer exposure (2 mins.). Panel C is an overnight exposure.

Treated and untreated WCEs were separated through a 10% SDS-PAGE under reducing conditions, transferred to a nitrocellulose membrane by electroblotting and probed with the anti-Pk mAb. Figure 36 shows that when WCEs from pT7ΔHNctPk and pT7ΔHNntPk transfected cells were treated with the enzyme PNGase F, the molecular weight of the minor ~68Kda band from the ntPk-tagged ΔHN expressed product was decreased by ~5Kda (lane 2) directly indicating that the ntPk-tagged glycoprotein had undergone N-linked glycosylation, albeit on only a small fraction of the expressed molecules. No reduction in mwt. was observed for the ct Pk-tagged ΔHN protein when incubated with PNGase F (lane 4). Furthermore, figure 22 shows that when WCEs from cells transfected with DNA encoding the truncated form of the Pk-tagged F are incubated with PNGase F, a decrease in the relative molecular weight is observed in the majority of the detectable molecules indicating the presence of N-linked oligosaccharides.

1.4.4 / Oligomeric structure of C-terminally Pk-tagged truncated SV5 F.

Chemical cross-linking analysis was performed as detailed in the methods section. Results from the chemical cross-linking analysis of C-terminally Pk-tagged truncated SV5 F are shown in figure 29 and demonstrated that, under non-reducing conditions, specific bands corresponding to monomers (~62Kda) and dimers (~124Kda) were produced. However, the gel was too smeared to be able to determine whether trimers (~188Kda) were formed. Increasing the concentration of the BS³ cross-linking reagent from 0-5mM influenced the molecular weight of the bands detected, ranging from predominantly monomers through to dimers and possibly trimers. The cross-linked extracts were also analysed under reducing conditions but no further evidence for the formation of trimers could be inferred.

1.5 / Summary.

Preliminary evidence suggests that:

- a) The presence of the Pk-tag at the termini of HN and F did not prevent their transient expression and subsequent detection with the anti-Pk mAb.

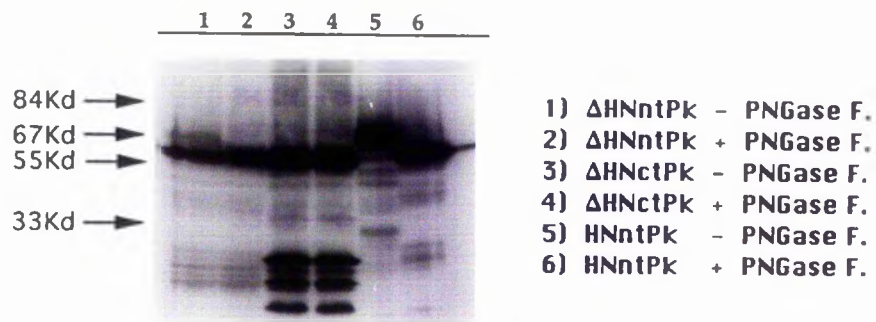


Figure 36: PNGase F analysis of Pk-tagged truncated SV5 HN (Δ HN) glycoproteins.

vTF7-3 infected 293 cells were transfected with the vectors; pT7 Δ HNntPk (lanes 1 & 2), pT7 Δ HNctPk (lanes 3 & 4) and pT7HNntPk (lanes 5 & 6). 16hrs post-transfection, WCEs were prepared in disruption buffer. Half of the samples were treated with the enzyme PNGase F, as detailed in the methods. Untreated/treated extracts were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb. Detection was by ECL.

- b) The Pk-tag was not apparently degraded, or clipped from the Pk-tagged HN & F proteins, by proteolytic activity from mammalian-cell proteases, or proteases encoded by the vTF7-3 vaccinia virus.
- c) Addition of the Pk-tag to the C-terminus of SV5 F did not prevent signal peptidase cleavage of the signal sequence to produce the type I glycoprotein topology.
- d) Addition of the Pk-tag to the C-terminus of the F glycoprotein did not prevent the post-translational proteolytic cleavage of F₀ to the disulphide-linked F₁ & F₂ subunits.
- e) The Pk-tagged HN and F glycoproteins were post-translationally modified by the addition of 6-10 Kda of N-linked oligosaccharides.
- f) The presence of the Pk-tag did not prevent the surface localisation and oligomerisation of the Pk-tagged HN & F glycoproteins.
- g) The presence of the Pk-tag did not prevent the recognition of the Pk-tagged HN & F glycoproteins by conformationally sensitive mAbs, which were raised against the native glycoproteins.
- h) Truncated forms of Pk-tagged HN & F were not secreted in a detectable form.
- i) Δ HNntPk was only N-linked glycosylated in a small proportion of the detected molecules. There was no evidence for N-linked glycosylation of the Δ HNctPk protein.
- j) The majority of Δ FctPk molecules were not in an oligomeric form, as determined by chemical cross-linking analysis.

SECTION 2 : ATTEMPTS AT PRODUCING CELL-LINES WHICH EXPRESSED Pk-TAGGED SV5 HN.

As demonstrated above, transient production and subsequent characterisation of Pk-tagged HN & F glycoproteins was achieved by utilising the vaccinia virus/T7-RDRP expression system. However, although this system is excellent for transient protein production for the initial characterisation studies, it can not be used for the production of stably expressing cell-lines with a view to the scale-up of protein production. For this

purpose, promoters which are recognised by endogenous cellular RNA polymerase II are utilised in plasmid vectors for the construction of stably expressing cell-lines.

2.1 / Attempts at the construction of cell-lines which *constitutively* expressed Pk-tagged SV5 HN.

The hCMV IE enhancer/promoter and RSV LTR are commonly used for driving gene expression in a variety of mammalian cells. Vectors containing these promoters, and sequences from Epstein-Barr virus for the episomal maintenance of the plasmid, were initially used in an attempt to produce stably expressing Pk-tagged HN cell-lines. For details on EBV-based vectors see chapter 3 in the introduction.

2.1.1 / Construction of EBV-based vectors which encode C - terminally Pk-tagged SV5 HN.

Two EBV-based vectors, pREP4 (RSV LTR) and pCEP4 (hCMV IE enhancer/promoter), were initially used to try and produce cell-lines which stably expressed C-terminally Pk-tagged HN. The plasmids pREP4HNctPk and pCEP4HNctPk (figure 37) were constructed by digestion of the plasmid pT7HNctPk with the REs Kpn1 and Hind3 with the insert containing the gene encoding the HNctPk glycoprotein being ligated into the similarly digested EBV-based vectors. The 5' and 3' regions of the sequence coding for Pk-tagged HN, within the plasmids pREP4HNctPk and pCEP4HNctPk, were sequenced to ensure fidelity of ligation.

2.1.2 / Transient protein production from the vectors *pREP4HNctPk* and *pCEP4HNctPk*.

293 and 293/E (constitutively express the Epstein-Barr virus EBNA-1) cells were transfected with the expression vectors pREP4HNctPk and pCEP4HNctPk, as detailed in the methods section. At 24, 36 and 48hrs post-transfection, the cells were processed for; Western blot analysis of WCEs, immunoprecipitation analysis of SCEs (using a pool of anti-HN mAbs (1b, 4b & 5d) & the anti-Pk mAb) or immunofluorescence analysis (again using anti-HN & anti-Pk mAbs). However, no protein production could be detected from either vector, transfected into either of the cell-lines, harvested at any of the times specified post-transfection, using the aforementioned means of detection (data not shown). Transfection

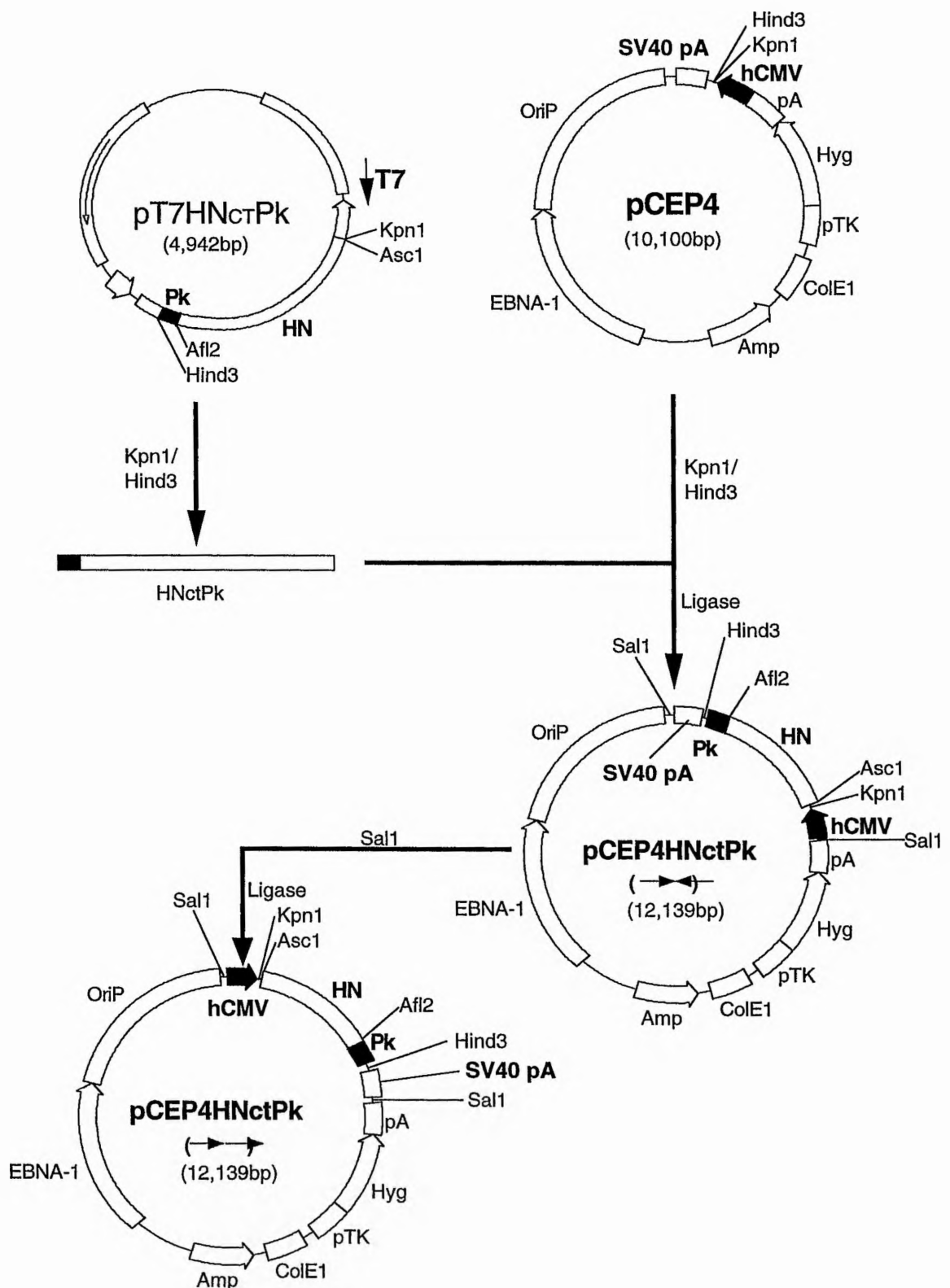


Figure 37: Construction of the EBV-based vectors which express the C-terminally Pk-tagged SV5 HN g'protein from the hCMV IE enhancer/promoter.

The plasmid pT7HNctPk was digested with the REs Kpn1 & Hind3, and the gene encoding the C-terminally Pk-tagged HN glycoprotein was ligated into the similarly digested EBV-based vector pCEP4, producing the vector - pCEP4HNctPk (-><-). pCEP4HNctPk (-><-) was digested with the RE Sal1, ligated and the vector pCEP4HNctPk (->->) was isolated in which the the HNctPk transcription unit was in the reverse orientation.

efficiency was monitored, in this experiment and throughout the work, by transfection of cells with the vector pCT7 β gal which encodes the β -galactosidase gene, where expression is driven by the hCMV major IE promoter. Transfection efficiencies were consistently in the range of 10-25% as determined by the β gal assay (details of this assay are given in the methods section).

2.1.3 /Attempts at the production of stably expressing cell-lines using the vectors pREP4HNctPk and pCEP4HNctPk.

Although no protein production could be detected in a transient expression analysis, drug-resistant cell-lines were produced as detailed in the methods section. This step was carried out in the hope that expression levels in a transient assay were too low to be readily detectable, and that if the majority of cells within a given population were expressing the Pk-tagged HN then detectable levels of protein may be produced. Briefly, 1×10^6 293 or 293/E cells were transfected with the above vectors. 48hrs post-transfection, the cells were passed into 12cm diameter petri dishes and the selective agent, hygromycin B, was added. 24 individual colonies were ring-cloned and expanded, with a master culture being prepared from the remaining colonies. Both the 24 individual cell-lines and the master cultures were analysed for protein expression using the techniques as described above. No production of Pk-tagged-HN was detected in any cell-line or master culture.

2.1.4 /Construction of the EBV-based vector encoding C - terminally Pk-tagged SV5 HN with a reversed transcription unit (figure 38).

Young *et al.* (1988) have observed that, at least for the EBV-based vector & gene of interest (goi) in use by them, the orientation of the goi transcription unit in relation to the transcription unit of the EBNA-1 gene was critical to the successful expression of the goi. When the goi transcription unit was in an opposite orientation to that of the EBNA-1 transcription unit then expression was negligible. However, upon reversing the goi transcription unit, high levels of expression of their goi was achieved. The HNctPk transcription unit within the vector pCEP4HNctPk was in this "unfavourable orientation". Therefore, the transcription unit of HNctPk was reversed to produce the more "favourable"

orientation. This was achieved by digesting the vector pCEP4HNctPk (now termed pCEP4HNctPk (-><-) to signify that the two transcription units are running in opposite directions) with Sal1 and religating the two fragments together where 50% of recombinants should have the transcription unit in the reversed direction. Such recombinants were determined by a diagnostic RE digestion. The vector in which the two transcription units are in the "favourable" orientation was designated pCEP4HNctPk (->->).

2.1.5 / Transient protein production from the vector pCEP4HNctPk (->->).

293/E cells were transfected with the expression vector pCEP4HNctPk (->->), as detailed in the methods section. At 24, 36 and 48hrs post-transfection, the cells were processed for; Western Blot analysis of WCEs, immunoprecipitation analysis of SCEs (using anti-HN & anti-Pk mAbs) or immunofluorescence analysis (again using anti-HN & anti-Pk mAbs). However, no protein production could be detected from WCEs from cells transfected with the more "favourable" vector, harvested at any of the times specified post-transfection, using the aforementioned means of detection (data not shown). Due to the lack of detecting protein expression, and the poor result from producing drug-resistant cell-lines with the previous vector when an initial transient result was negative, the isolation of drug-resistant cell-lines from cells transfected with the vector pCEP4HNctPk (->->) was not attempted.

2.2 / Construction of transient expression vectors which produce Pk-tagged SV5 HN via an hCMV IE promoter.

Due to the inability to detect recombinant protein production from the EBV-based vectors described above, an alternative vector, pcDNA1/amp (figure 38), was employed to try and detect protein production driven by the hCMV IE enhancer/promoter in a transient assay. This vector had the advantage over the EBV-based vectors in that downstream of the hCMV IE enhancer/promoter was the T7 ϕ 10 promoter which could be used as a control for Pk-tagged HN expression. Furthermore, pcDNA1/amp also possesses the simian virus 40 origin of replication, which should result in the amplification of the plasmid when transfected into Cos cells (which express the SV40 large T antigen) with a subsequent increase in protein production. Furthermore, n-terminally Pk-tagged HN was also tested for

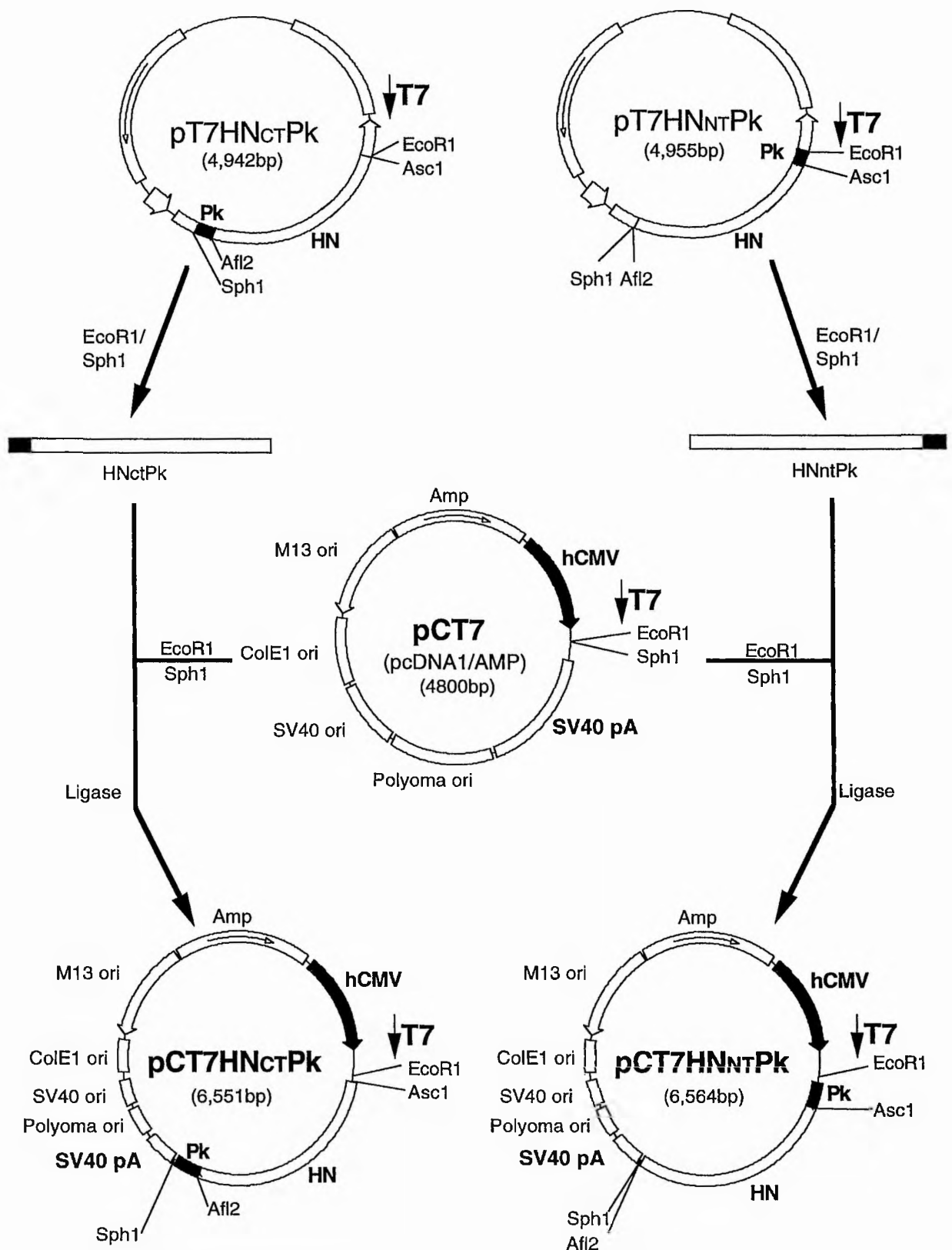


Figure 38: Construction of the vectors which express the Pk-tagged SV5 HN g'proteins from the hCMV IE enhancer/promoter.

The plasmids pT7HNctPk & pT7HNntPk were digested with the REs EcoR1 & Sph1, and the genes encoding the Pk-tagged HN glycoproteins were ligated into the similarly digested vector pCT7 (pcDNA1/amp), producing the vectors - pCT7HNntPk & pCT7HNctPk.

expression under the hCMV IE promoter since a lack of expression from the previous experiments may be due to an inherent property of the c-terminally Pk-tagged HN.

2.2.1 / Construction of pcDNA-based vectors for the transient production of N & C-terminally Pk-tagged SV5 HN (figure 38).

The expression vectors pCT7HNctPk and pCT7HNntPk were constructed by digestion of the plasmids pT7HNctPk & pT7HNntPk with the restriction endonucleases EcoR1 & Sph1. The inserts containing the genes encoding the HNctPk and HNntPk glycoproteins were ligated into the similarly digested plasmid pcDNA1/ amp. Positive recombinants were screened by a diagnostic restriction endonuclease digestion analysis.

2.2.2 / Production of N & C-terminally Pk-tagged SV5 HN from the pcDNA-based vectors.

To test whether the pcDNA-based constructs expressed Pk-tagged-HN glycoproteins, expression was initially driven from the T7 ϕ 10 promoter, in combination with the recombinant vaccinia virus encoding the T7-RDRP. Figure 39 shows the result of analysing WCEs from transfected 293 cells which were separated through an SDS-PAGE, electroblotted and probed with the anti-Pk mAb. Bands of ~70 Kda were easily detected from cells transfected with two different clones of the pCT7HNctPk & pCT7HNntPk constructs.

Since the Pk-tagged-HN glycoproteins could be detected when expression was driven from the T7 ϕ 10 promoter, the ability to detect tagged-HN when expression was driven by the hCMV IE enhancer/promoter, when transfected into 293 or Cos-7 cells, was investigated. The transfections were performed in combination (in a 1:1 ratio) with a plasmid (pMHVA) expressing the adenovirus VAI product, (with or without 2-aminopurine, 2-ap, added 32hrs post-transfection) or with the plasmid pUC19 (for details of the influence of VAI and 2-ap on transcription/translation, see chapter 4 in the introduction). Figure 40 shows the result of preparing WCEs 48hrs post-transfection and processing as for the T7 driven analysis above. No bands could be detected even after an overnight ECL exposure.

Furthermore, since proteolytic cleavage of the Pk-tag epitope maybe occurring and thus preventing detection of expressed protein by Western blot analysis using the anti-Pk mAb, an indirect immunofluorescence analysis using anti-Pk or a pool of anti-HN mAbs (1b,

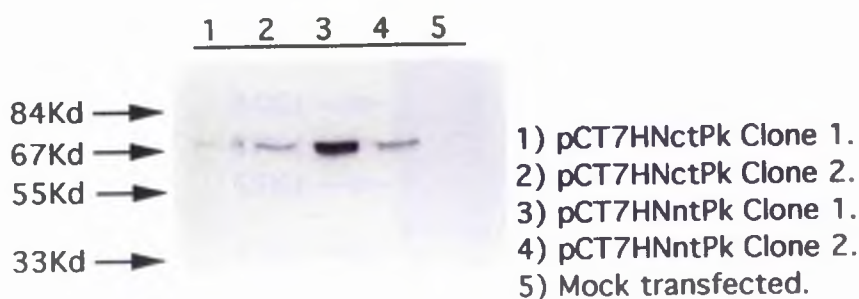


Figure 39: Transient expression of Pk-tagged HN glycoproteins from the pcDNA-based vectors under control of the T7 promoter.

vTF7-3 infected 293 cells were transfected with 2 clones of the plasmids pCT7HNctPk (lanes 1 & 2) & pCT7HNntPk (lanes 3 & 4), or were mock transfected (lane 5). 16hrs post-transfection WCEs were prepared in x1 disruption buffer and were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by detection using ECL. Clone 2 of pCT7HNctPk and clone 1 of pCT7HNntPk were used for further analysis (see below).

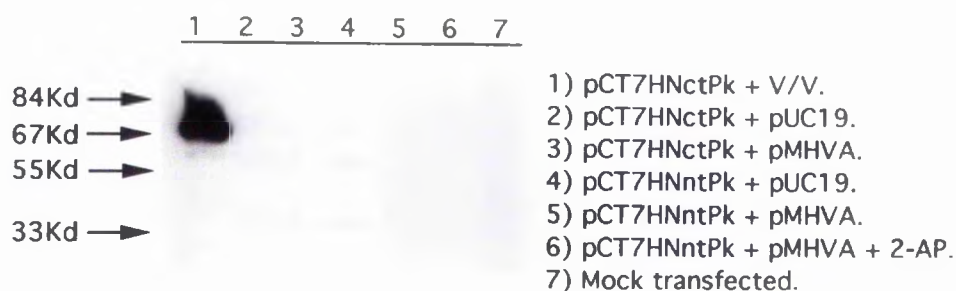


Figure 40: Transient expression of Pk-tagged HN glycoproteins from the pcDNA-based vectors under control of the hCMV major IE enhancer/promoter.

293 cells were transfected with the plasmids pCT7HNctPk (lanes 2 & 3) & pCT7HNntPk (lanes 4-6) in combination with a 1:1 ratio of the adenovirus VAI-expressing plasmid pMHVA (lanes 3, 5 & 6) or pUC19 (lanes 2 & 4). 32hrs post-transfection 2-aminopurine (2-AP) was added to a final concentration of 10mM to one set of transfections (lane 6). 48hrs post-transfection WCEs were prepared in x1 disruption buffer and were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by detection using ECL (exposure was overnight). As a positive control, an extract of T7 expressed (via vaccinia virus, V/V) HNctPk was included.

4b & 5d) was carried out. Unfortunately, no expression of HN was detected in transfected Cos-7 or 293 cells, even with the anti-HN mAbs (data not shown).

To ascertain whether the hCMV IE enhancer/promoter was functional, the gene encoding the enzyme β -galactosidase was cloned into the vector pCT7HNctPk, thereby replacing the HNctPk gene. This vector, pCT7 β gal, was transfected into 293 or Cos-7 cells and 48hrs post-transfection the cells were stained for evidence of β -galactosidase activity, as detailed in materials and methods. Between 20-30% of the cells stained blue indicating a functional hCMV IE enhancer/promoter element (data not shown).

2.3 / Attempts at the construction of cell-lines which inducibly expressed Pk-tagged SV5 HN.

Due to the inability to detect the production of Pk-tagged HN when transcription was driven by the hCMV IE enhancer/promoter or the RSV LTR, an alternative, inducible, expression system first developed by Gossen & Bujard based upon the tetracycline transactivator, was employed (for details of this system see chapter 3 in the introduction). This approach was chosen due to the above failures when expression was driven from a constitutive promoter and from findings from other groups (R.A. Lamb, personal communication) in which stable cell-lines which constitutively express SV5 HN could not be isolated. Therefore, this may indicate that the SV5 HN protein is toxic to cells if expressed constitutively.

2.3.1 /Construction of pTET vectors which encode N & C - terminally Pk-tagged SV5 HN.

The plasmids pTETHNctPk and pTETHNntPk (figure 41) which encode Pk-tagged HN under the control of the tetracycline responsive promoter (TRP) were constructed by digestion of the plasmids pCT7HNctPk & pCT7HNntPk with the restriction endonucleases EcoR1 and Xba1. The inserts containing the genes encoding the HNctPk or HNntPk glycoproteins were ligated into the similarly digested pTET vector (pUHD 10-3). The 5' region of the sequence encoding the Pk-tagged HN, within the plasmids pTETHNctPk and pTETHNntPk, was sequenced to ensure fidelity of ligation.

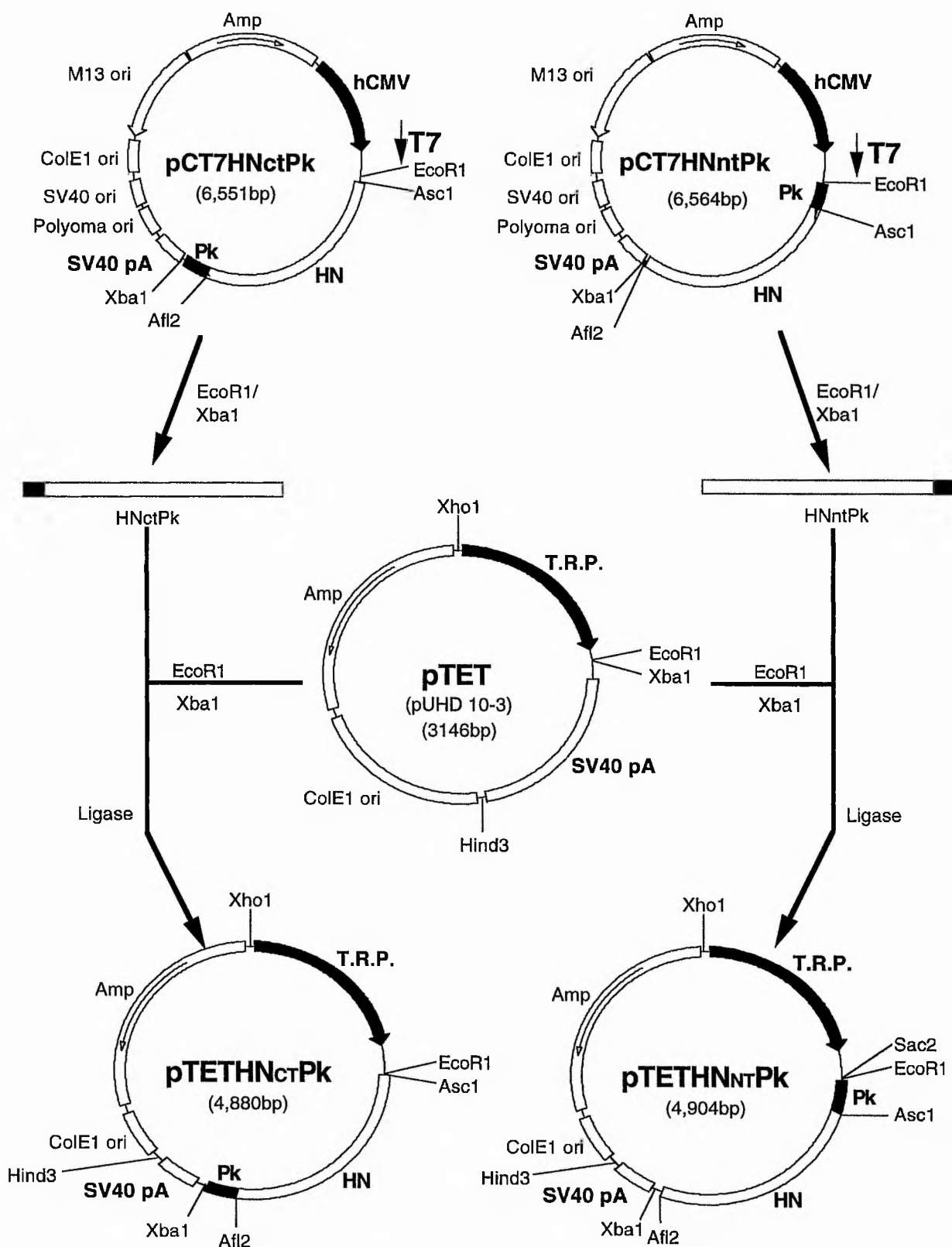


Figure 41: Construction of the vectors which inducibly express Pk-tagged SV5 HN glycoproteins.

The plasmids pCT7HNctPk & pCT7HNntPk were digested with the REs EcoR1 & Xba1, and the genes encoding the Pk-tagged HN g'proteins were ligated into the similarly digested vector pTET, producing the tetracycline-responsive vectors - pTETHNntPk & pTETHNctPk.

2.3.2 / Transient protein production from the vectors pTETHNctPk and pTETHNntPk.

293/E cells were transfected with the expression vectors pTETHNctPk and pTETHNntPk and the driver plasmid pMR-tTa/neo (in a 1:1 ratio), in the absence of tetracycline, as detailed in the methods section. The plasmid pMR-tTa/neo, figure 44, expresses the tTa transactivator as an in-frame fusion protein with the neo^r gene. The two genes are separated by a sequence coding for the foot & mouth disease virus 2A cleavage site which is cleaved, co-translationally, resulting in the separation of the fusion protein into the two active gene products: tTa & neo^r (Precious *et al.*, 1995). At 48hrs post-transfection, the cells were processed for Western blot analysis of WCEs and immunoprecipitation analysis of SCEs (using the anti-Pk mAb). Figure 43 shows that a band of ~70Kda was clearly detectable from extracts of cells transfected with the vector encoding the N-terminally Pk-tagged HN glycoprotein. However, no band was detected from extracts of cells transfected with the plasmid encoding the C-terminally Pk-tagged HN glycoprotein.

2.3.3 / Construction of pTEP4 vectors which encode N & C - terminally Pk-tagged SV5 HN.

The pTET vectors do not possess drug-resistant markers and are not stably retained as episomes in the nucleus of cells. Therefore, the pTETHNct & pTETHNntPk vectors were altered by the addition of the EBV-based backbone from the pCEP4 series of vectors. The plasmids pTETHNctPk & pTETHNntPk were digested with the RE Hind3 and a dsDNA oligonucleotide containing an Xho1 RE site and Hind3 cohesive-ends was ligated into the digested vectors to produce the plasmids, pTETHNctPk/X2 and pTETHNntPk/X2. These plasmids were then digested with the RE Xho1, with the inserts containing the transcription unit encompassing the TRP/HNct or ntPk/SV40 pA being ligated into the EBV-backbone of the vector pCEP4HNctPk (->->), which had been previously digested with the RE Sal1. The orientation of the inserted transcription unit in relation to the direction of transcription of the EBNA-1 gene was determined by a diagnostic restriction endonuclease digestion. Plasmids of both orientation were chosen for further analysis and were designated: pTEP4HNctPk (->->) & (->-<) and pTEP4HNntPk (->->) & (->-<), (figure 42).

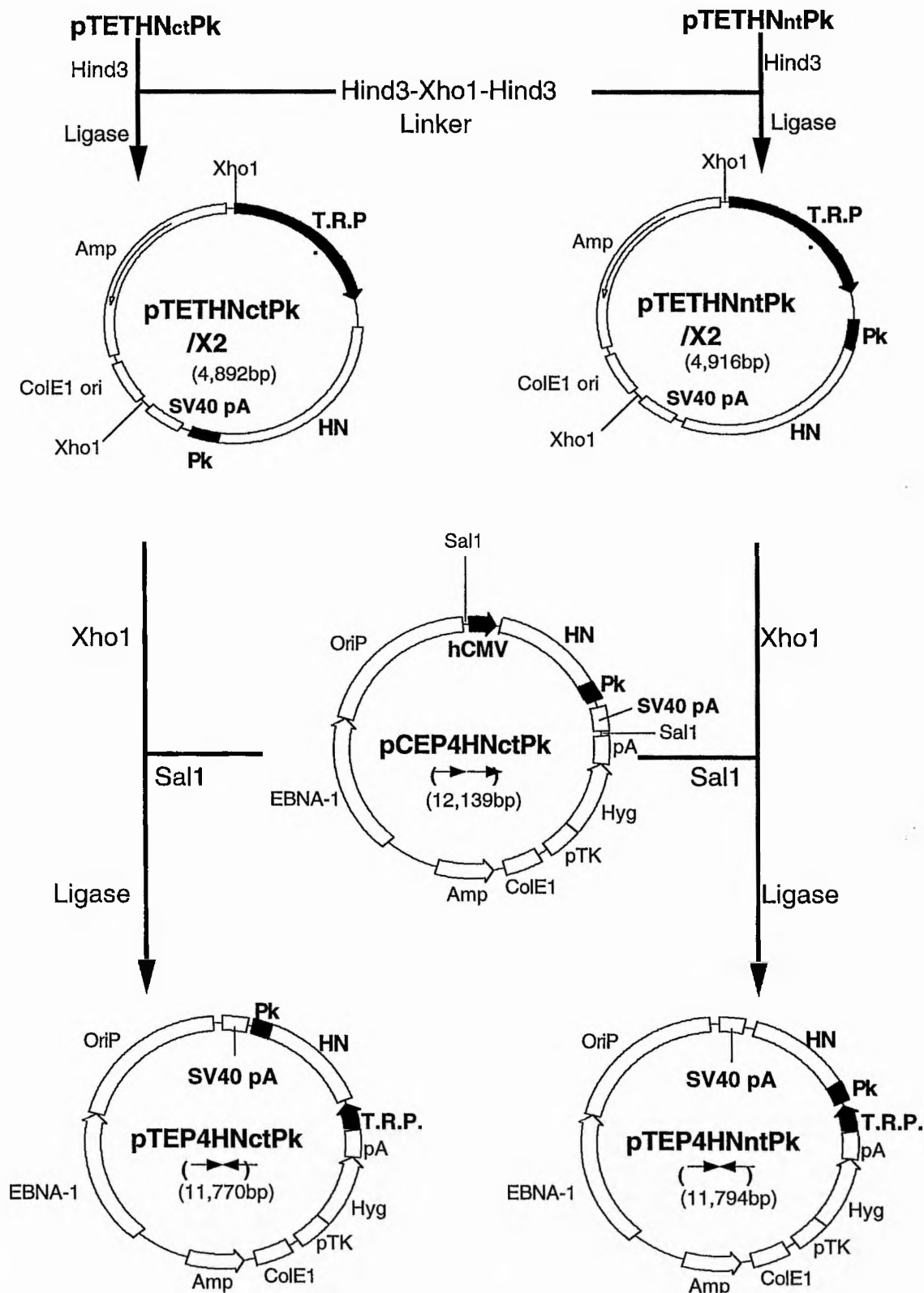


Figure 42: Construction of EBV-based vectors for the inducible expression of Pk-tagged SV5 HN glycoproteins.

A second Xho1 RE site was incorporated into the plasmids pTETHNctPk & pTETHNntPk via a dsDNA oligonucleotide linker which was ligated into the Hind3 digested vectors. The TRP/HN/pA transcription units were excised from the plasmids by a Xho1 digestion and cloned into the EBV-backbone of the Sal1 digested vector pCEP4HNctPk (->->) to produce the vectors pTEPHNntPk (->->) & (-><-) and pTEPHNctPk (->->) & (-><-). Orientations were determined by a diagnostic RE digestion. (->->) orientations are not shown.

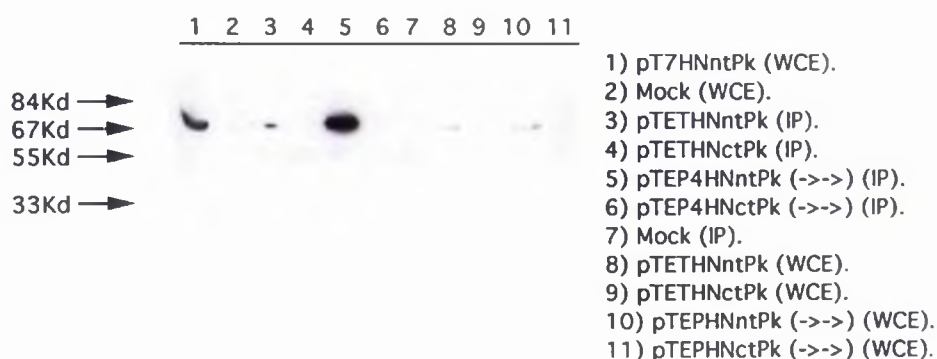


Figure 43: Transient expression analysis of Pk-tagged SV5 HN expressed from a tetracycline responsive promoter.

293/E cells were transfected with the plasmids; pTETHNntPk (lanes 3 & 8), pTETHNctPk (lanes 4 & 9), pTEP4HNntPk (->->) (lanes 5 & 10), pTEP4HNctPk (->->) (lanes 6 & 11), in a 1:1 ratio with the tTa driver plasmid pMR-tTa/neo, or were mock transfected (lanes 7 & 12). 48 hrs post-transfection, WCEs were prepared in disruption buffer from which an aliquot was adjusted to x1 disruption buffer. The remainder of the extract was centrifuged and the resulting SCEs was immunoprecipitated (IP) using the anti-Pk mAb together with a pool of anti-HN mAbs (1b, 4b & 5d). Samples of both the WCEs (lanes 8-12) and the IPs (lanes 3-7) were electrophoresed through a 10% SDS-PAGE, electroblotted and probed with the anti-Pk mAb. Detection was by ECL. A WCE from vTF7-3 infected 293 cells transfected with the plasmid pT7HNctPk, was included as a positive control (lane 1).

2.3.4 / Transient protein production from the vectors pTEP4HNctPk (->->) and pTEP4HNntPk (->->).

293/E cells were transfected with the expression vectors pTEP4HNctPk (->->) & pTEP4HNntPk (->->) and the driver plasmid pMR-tTa/neo (in a 1:1 ratio) in the absence of tetracycline, as detailed in the methods section. At 48hrs post-transfection, the cells were processed for Western blot analysis of WCEs and immunoprecipitation analysis of SCEs (using the anti-Pk mAb). Figure 43 shows the result of separating the WCE or immunoprecipitate through a 10% SDS-PAGE, electroblotting and probing with the anti-Pk mAb. A band of ~70Kda is clearly detectable from extracts of cells transfected with the vector encoding the N-terminally Pk-tagged HN glycoprotein. However, as with the pTETHNctPk vector, no band was detected from extracts of cells transfected with the EBV-based plasmid encoding the C-terminally Pk-tagged HN glycoprotein. Immunofluorescence analysis in 293 cells, using a pool of anti-HN mAbs (1b, 4b & 5d), also failed to detect any expression from the pTEP4HNctPk (->->) construct (data not shown).

2.3.5 / Attempt at the establishment of stable cell-lines which inducibly expressed N-terminally Pk-tagged SV5 HN using the vectors pEBV-tTa/neo & pTEP4HNntPk (->->).

1×10^6 293 cells were transfected with the EBV-based vectors pEBV-tTa/neo (a modified version of the vector pMR-tTa/neo, in which EBV sequences have been added to the vector for episomal maintenance, B. Precious Univ. of St. Andrews) and pTEP4HNntPk (->->) in a 1:1 ratio in the presence of $3 \mu\text{gml}^{-1}$ tetracycline. Forty eight hours post-transfection, the cells were passed into 12cm diameter petri dishes and the selective agents, G418 and hygromycin B were added, again with tetracycline present. After 4 weeks in the selective medium only one colony was produced. This single colony was analysed for expression of HNntPk and tTa by immunofluorescence analysis using anti-VP16 anti-serum (recognises the acidic domain of VP16 from Herpes simplex virus, which comprises one half of the tTa fusion protein. For more detail on tTa see chapter 4 in the introduction) and anti-HN (1b, 4b & 5d) & anti-Pk mAbs. However, no expression could be detected for either proteins (data not shown).

Due to only one drug-resistant colony being produced from the transfection/selection procedure, two approaches to try and increase the number of colonies produced, and therefore the probability of isolating HNntPk expressing cells, were adopted.

(1) 293 cells which constitutively express the EBV EBNA-1 protein (293/Es) were utilised to try and increase the number of drug-resistant colonies produced. [Cell-lines which express EBNA-1 appear to better facilitate the isolation of stable cell-lines when employing EBV-based vectors. This is most probably due to the initial presence of the EBNA-1 protein in the nucleus, which aids in the replication of the incoming EBV-based vectors and thus of their stable maintenance]. Unfortunately, these 293/E cells expressed the neo^r gene product and are thus G418 resistant. Therefore, the various tTa/neo constructs developed in our laboratory could not be employed, and new vectors possessing the gene encoding hygromycin B phosphotransferase (tTa/hygB) were constructed, as detailed in section 2.3.6. Furthermore, since the inducibly expressing HNntPk EBV-based plasmids encode a gene for resistance to hygB then new EBV-based vectors were constructed in which resistance was conferred by the histidinol dehydrogenase gene, see section 2.3.7(the construction of these particular vectors was also necessary for experiments performed in section 3 in which the presence of an SV40 *ori* was required).

(2) The production of a 293-based cell-line which constitutively expressed high levels of the tTa protein from the EBV-based vector pEBV-tTa/neo was attempted (as detailed in section 2.3.8). This should render the 293 cells EBNA-1 expressing and should thus aid in the isolation of drug-resistant colonies, as described above. Furthermore, since the cells will already be constitutively expressing the tTa protein, this will result in only the one large EBV-based plasmid (pTEP4HNntPk (->->)-hyg^r, or pTEP8HNntPk (->->)-his^r (will be described in section 2.3.7)) being transfected, thus increasing the transfection efficiency.

2.3.6 /Construction of, and expression from, vectors which encode tTa/hygB.

The gene encoding hygB was PCR amplified from the vector pCEP4 together with the thymidine kinase pA sequence using primers which had Apa1 RE sites incorporated. The PCR product was subsequently digested with Apa1 and ligated into the similarly digested

vector pMR101, thereby replacing the neo^r gene, SV40pA and polyoma origin of replication to produce the plasmid pMR-2A/hygB (figure 44). The gene encoding the tTa fusion protein was excised from the plasmid pMR-tTa/neo using the REs BamH1 & Afl2 and ligated into the similarly digested plasmid pMR-2A/hygB to produce the vector pMR-tTa/hygB (figure 44). This vector was tested in a transient expression assay together with the vector pTEP4HNntPk (->->) in 293/E cells. N-terminally Pk-tagged HN could be detected from a WCE used in a Western blot analysis when probed with the anti-Pk mAb (data not shown).

As with the vector pMR-tTa/neo, pMR-tTa/hygB has to integrate into the cellular genome before drug-resistant cells can be produced. Therefore, pMR-tTa/hygB was modified by the addition of EBV sequences for the episomal maintenance of the vector in primate-cell nuclei. Thus, pMR-tTa/hygB was digested with the REs Spe1 & Sfi1 and the tTa/hygB transcription unit was ligated into the plasmid pREP8 which had been digested with the REs Xba1 & Sfi1, producing the vector pEBV-tTa/hygB (figure 45). This vector was also tested in a transient expression assay and shown to express functional tTa (figure 50).

For future experiments (see section 3), a driver plasmid without the SV40 origin of replication and which can only integrate into the cellular genome was required. Therefore, pMR-tTa/hygB was digested with the REs Spe1 & Sfi1 and the tTa/hygB transcription unit was ligated into the plasmid pGEM-9zf(-) which had been digested with the REs Spe1 & Sfi1, producing the vector pGEM-tTa/hygB (figure 45). This vector was also tested in an expression assay as detailed below, and shown to express functional tTa (figure 46).

Furthermore, to ensure that the hygB^r gene product was expressed and functional, stable cell lines expressing β -galactosidase were constructed as follows. 1×10^6 293/E cells were transfected, in the presence of tetracycline, with the vectors pGEM-tTa/hygB & pTET β gal (pUHD-15) in a 1:1 ratio. 48 hrs post-transfection hygB was added to the transfected cells and to naive cells. All the cells in the control plate were dead within 1 week, and drug-resistant colonies grew within 2 weeks from the transfected pool of cells (thereby demonstrating the presence of a functional hygB^r gene product). These colonies were mixed together to form a drug-resistant master culture. The cells were seeded into two 35mm petri dishes, with tetracycline either absent or present. Cells were fixed 36hrs later and stained for

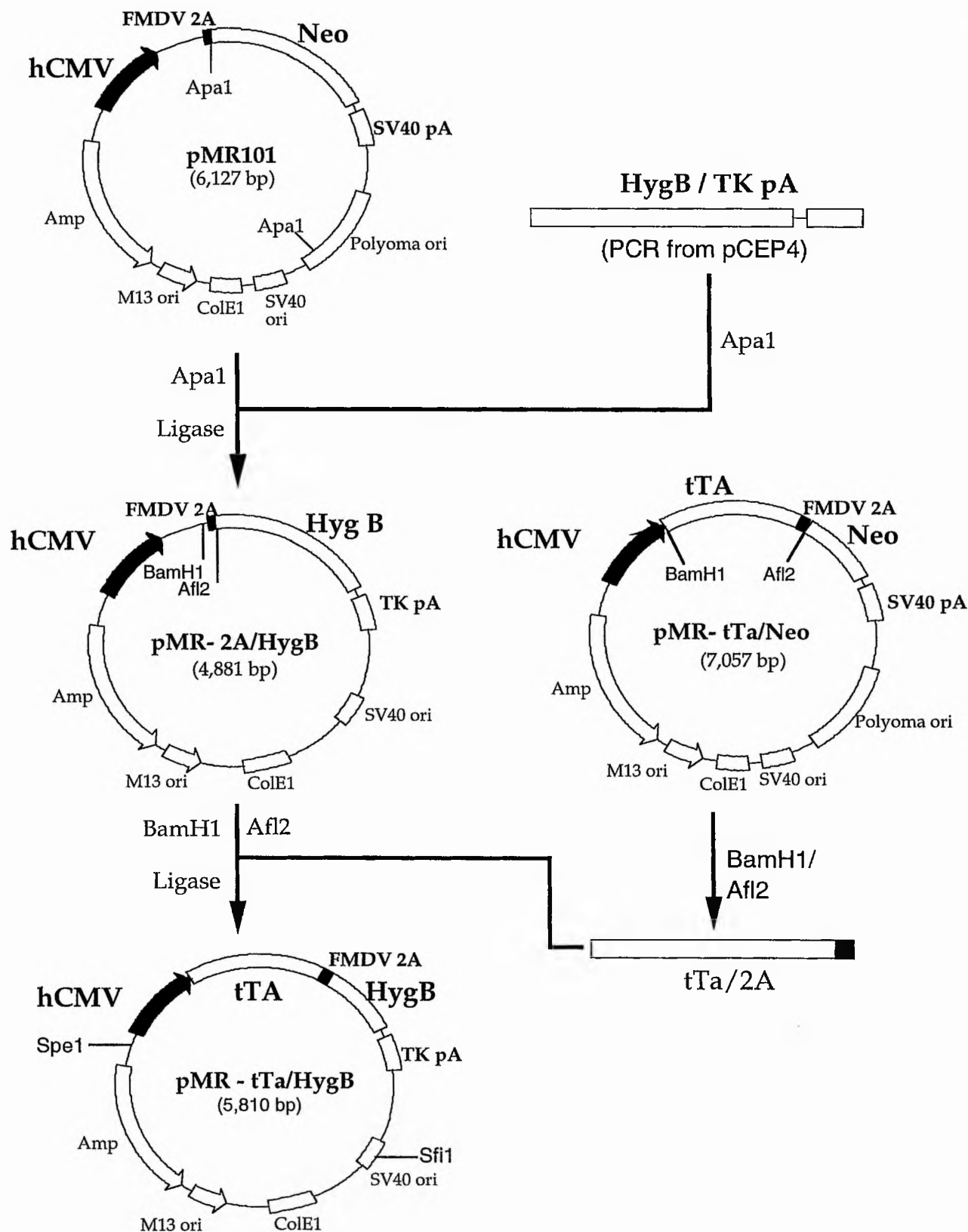


Figure 44: Construction of the vector for the co-translation of tTa and *hygB*.

The open reading frame of *hygB* and the TK pA were PCR amplified from the vector pCEP4 using primers which contained Apa1 RE sites. The PCR product was digested with Apa1 and subsequently ligated into the similarly digested vector pMR101. The DNA sequence encoding tTa/2A was excised from the vector pMR-tTa/ neo using the REs BamH1 & Afl2 and ligated into the similarly digested plasmid, pMR-2A/hygB to produce the vector pMR-tTa/hygB.

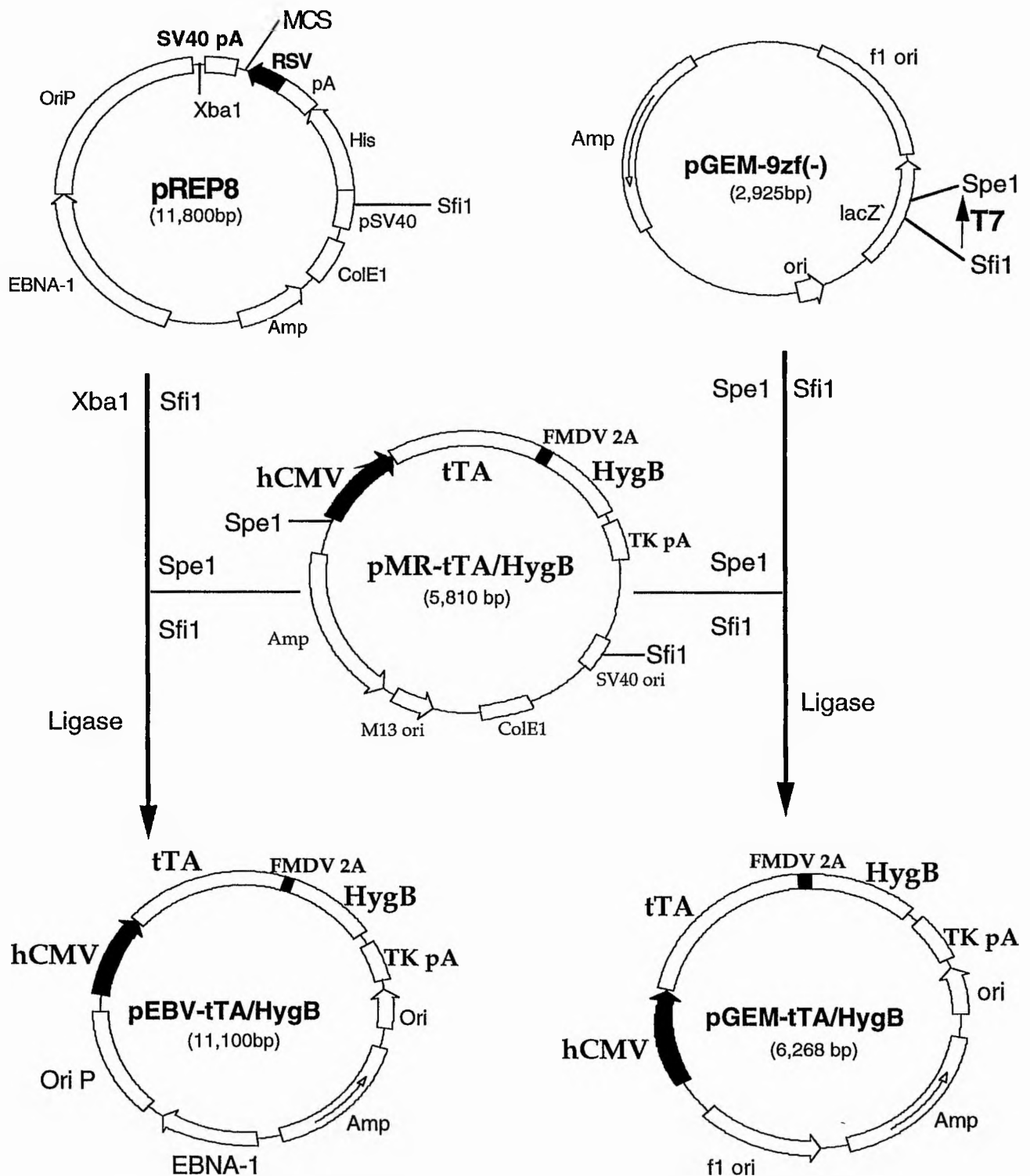
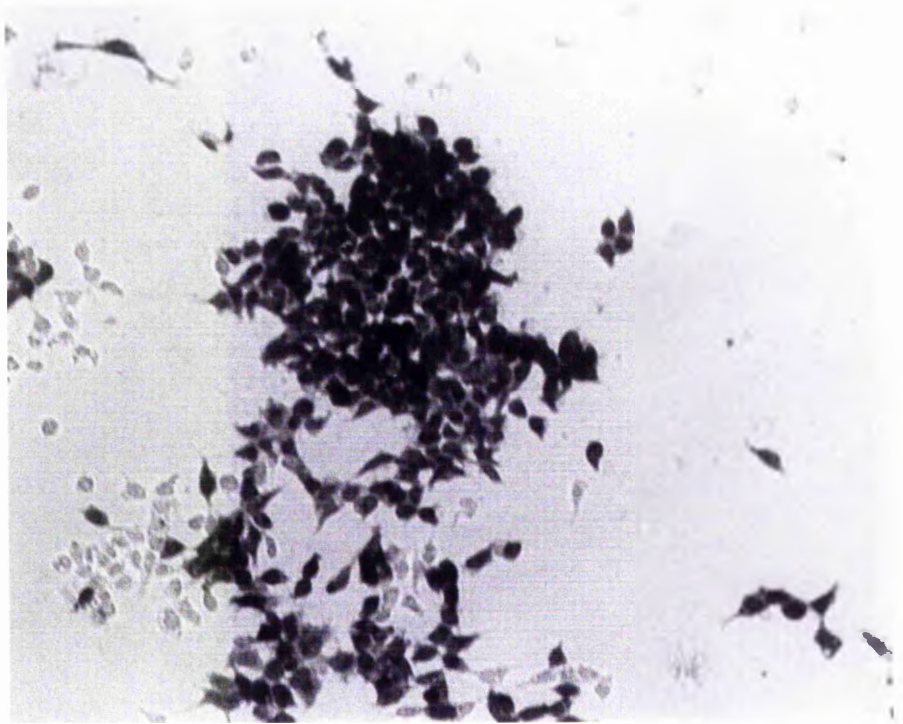


Figure 45: Construction of pEBV-tTa/hygB and pGEM-tTa/hygB vectors.

The DNA sequence containing the transcription unit encoding the tTa/hygB fusion protein was excised from the plasmid pMR-tTa/hygB using the REs Spe1 & Sfi1, and ligated into the similarly digested plasmid pGEM-9zf(-) to produce the vector pGEM-tTa/hygB. The EBV-based vector pREP8 was digested with the REs Xba1 & Sfi1 and the insert excised from pMR-tTa/hygB was ligated into the plasmid to produce the vector pEBV-tTa/hygB.

A)



B)

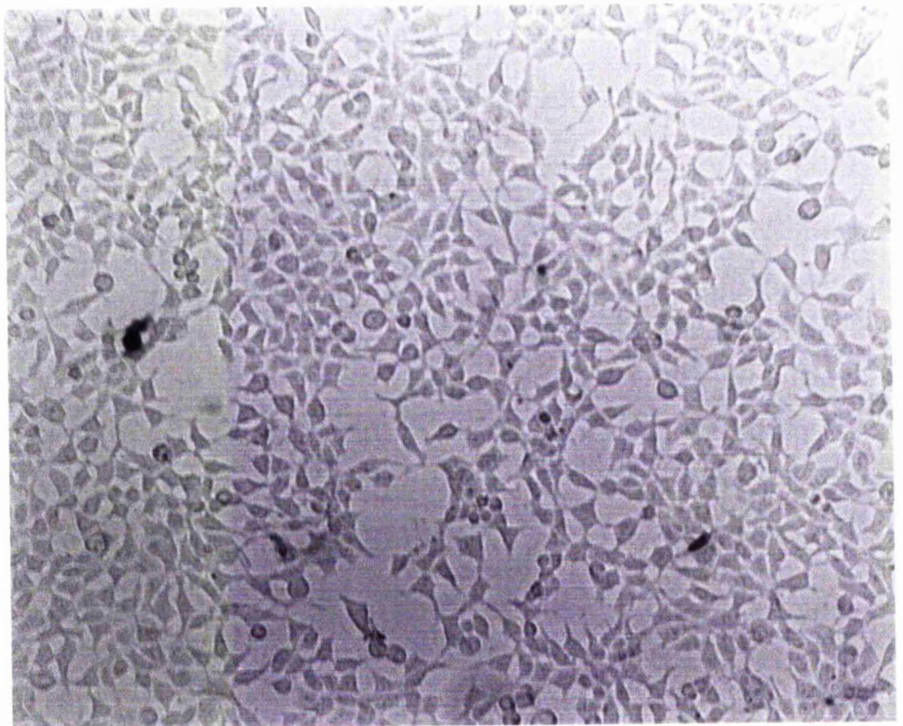


Figure 46: Functional analysis of tTa & hygromycin B phosphotransferase (hygB PT).

293 cells were transfected with the plasmids pGEM-tTa/hygB and pTRP β gal in the presence of tetracycline (3 μ g/ml). 48 hrs post-transfection, transfected cells, and an untransfected control population, were incubated in medium containing hygromycin B (and tetracycline) at a final concentration of 250 μ g/ml. Within two weeks of incubation in selective medium all the control cells had died. Approximately 50 colonies grew from the transfected cell population, thereby demonstrating the production of functional hygB PT. The colonies were passed into two 35 mm petri dishes, in the presence (panel B) or absence (panel A) of tetracycline. 48 hrs post-passing the cells were fixed and stained for detection of β -galactosidase, as detailed in the methods section.

β -gal activity as detailed in materials and methods. Figure 46 clearly shows the expression of β -gal in cells grown in the absence of tet^C, with very little β -galactosidase activity being evident from the cells grown in the presence of 3 μgml^{-1} tet^C.

2.3.7 / Construction of, and expression from, the EBV-based vectors which inducibly expressed the HNntPk glycoprotein and confer resistance to histidinol (figure 47).

The vectors pTEP8HNntPk (->->) & (-><-) were constructed as follows. The transcription unit encompassing the hCMV IE promoter/P gene/pA was excised from the plasmids pCEP8/P (->->) & (-><-) using the RE Sal1. The transcription unit comprising the TRP/HNntPk/pA was excised from the plasmid pTETHNntPk using the RE Xho1 and ligated into the EBV-backbone of the Sal1 digested pCEP8/P (->->) & (-><-) vectors to produce the desired plasmids pTEP8HNntPk (->->) & (-><-). pTEP8HNntPk (->->) was shown to express HNntPk, initially via an immunofluorescence analysis using anti-Pk mAb (figure 48). [A comparison of the levels of HNntPk protein produced when expressed from the two vectors was performed via a transient expression assay, see section 2.4].

2.3.8 /Construction of cell-lines which constitutively expressed tTa from an episomally maintained EBV-based vector.

1x10⁶ 293 cells were transfected with 3 μg s of the vector pEBV-tTa/neo. 48hrs post-transfection the cells were passed into a 25cm² flat and geneticin was added to a concentration of 400 μgml^{-1} . Resistant cells had begun to grow approximately 1 week later when the drug-resistant population were pooled together and seeded into individual wells of a 96-well dish, at ~ 10-20 cells per well. Individual wells were then screened for tTa expression by immunofluorescence using the anti-VP16 anti-serum. The clone showing the most number of cells expressing tTa at the highest level was then passed into a 12cm diameter petri and individual colonies (10 in total) were ring-cloned and screened for tTa expression as detailed above. 4 clones (\neq 1, \neq 2, \neq 3 & \neq 6) showing variable levels of tTa expression (figure 49) were further analysed. To confirm that the tTa produced is of the correct mwt., and to compare the level of tTa expression between the four clones, a Western blot analysis, was performed. WCEs from the four tTa expressing cell-lines were prepared in 1% NP40 and their protein content was assayed using the BCA method. The WCE was

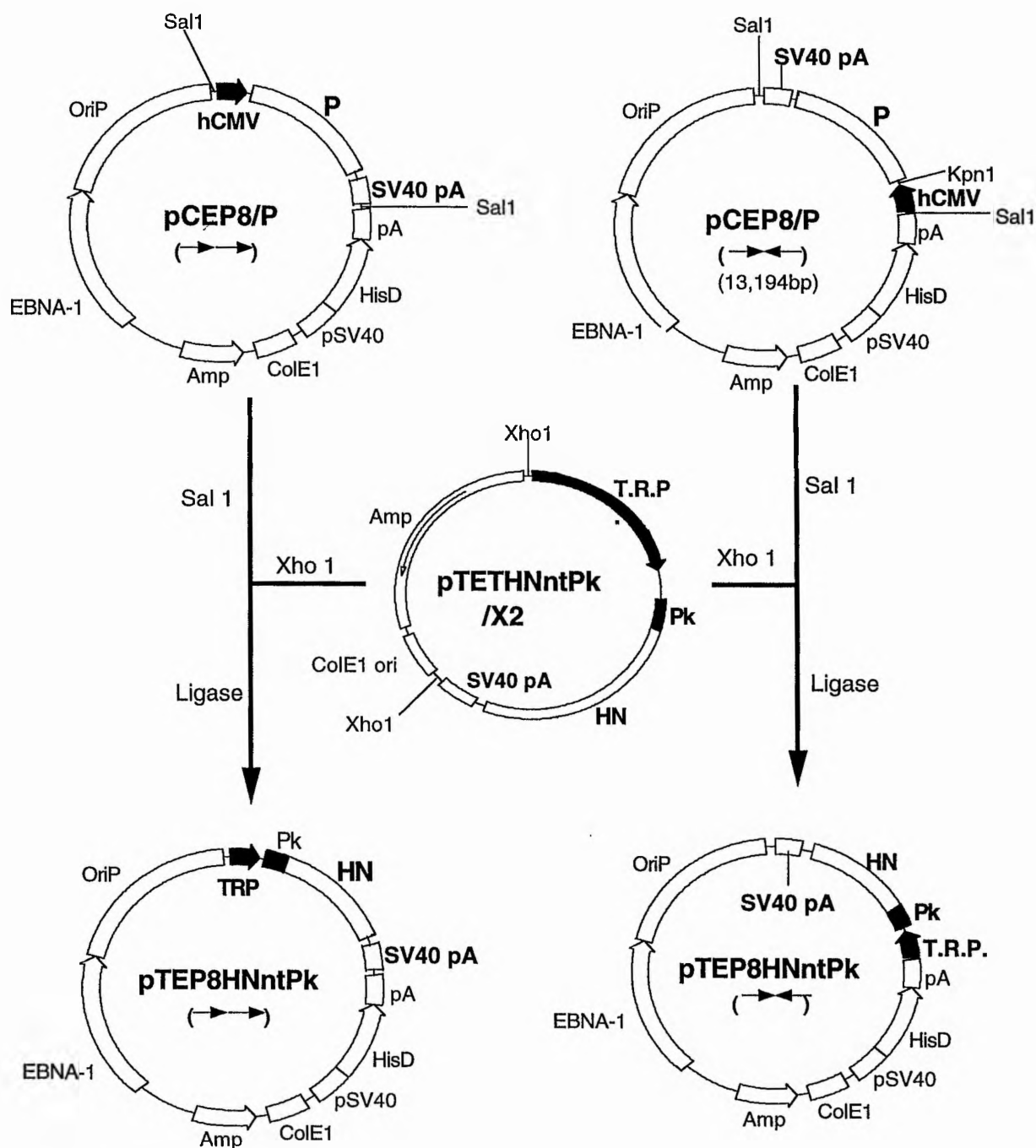


Figure 47: Construction of LTag-amplifiable EBV-based vectors which inducibly expressed N-terminally Pk-tagged SV5 HN.

The plasmids pCEP8/P (->->) & (-><-) were digested with the RE Sal1 and the transcription unit TRP/HNntPk/pA was excised, from the plasmid pTETHNntPk/X2 via a Xho 1 digestion, and ligated into the Sal 1 digested vectors.

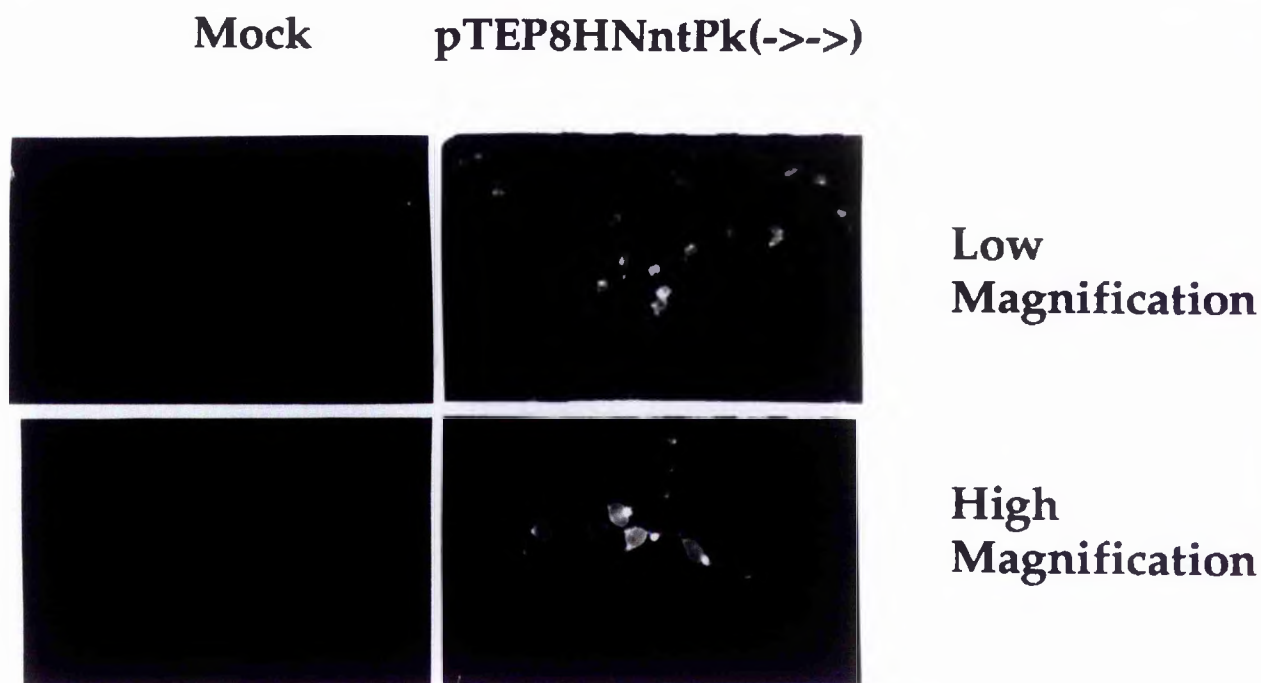


Figure 48: Immunofluorescence analysis of pTEP8HNntPk (->->) transfected 293/E cells.

293/E cells were transfected with the vectors pTEP8HNntPk (->->) & pGEM-tTA/hygB (1:1 ratio) or pCEP4 & pGEM-tTA/hygB (1:1 ratio, mock transfection). 48hrs post-transfection, cells were fixed & permeabilised then incubated with the anti-Pk mAb followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant. Low magnification x10. High magnification x40.

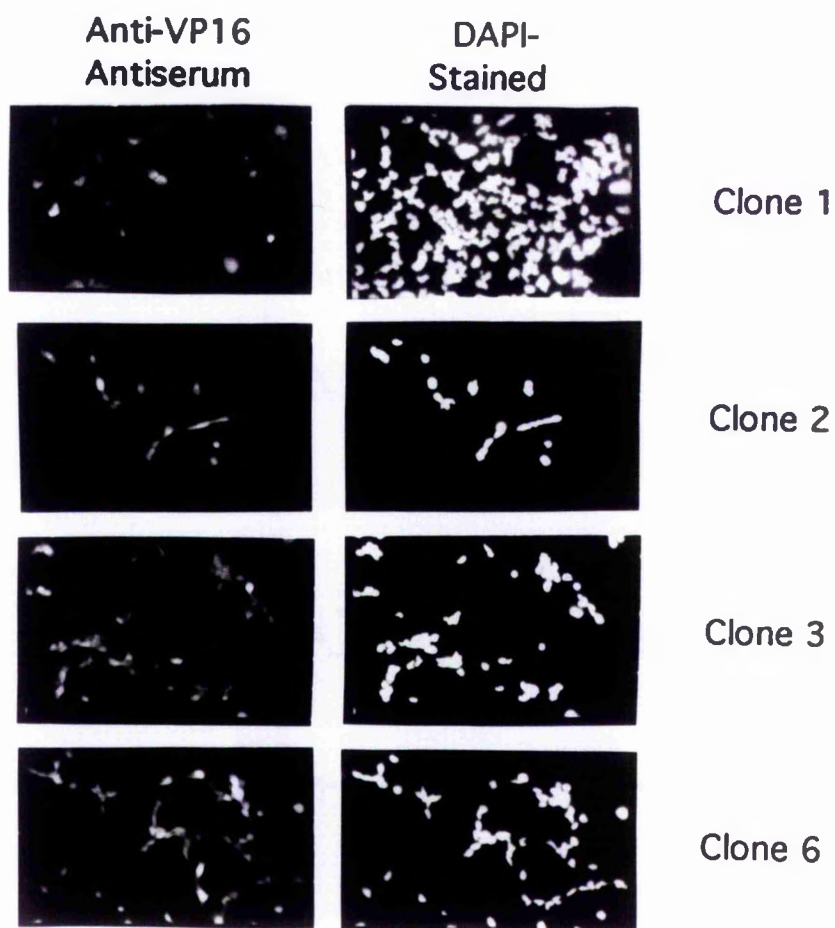


Figure 49: Immunofluorescence analysis of 4 isolated 293-based cell-lines transfected with an EBV-based plasmid producing the tTa protein.

293 cells were transfected with the vector (pEBV-tTa/neo). Drug-resistant cell-lines were isolated by selection in geneticin for two weeks (see text for details). Four clones were tested for tTa production by an immunofluorescence analysis using anti-serum to the acidic domain of herpes simplex virus virion protein 16 (VP16: one half of the tTa fusion protein).

adjusted to x1 disruption buffer and equivalent amounts of protein, from the four cell-lines, were separated through a 10% SDS-PAGE, electroblotted and probed antiserum to the acidic domain of virion protein 16 (VP16) from herpes simplex virus. Figure 50 shows that bands of ~ 40Kda, corresponding to the predicted mwt. of full length tTa could be detected from the WCEs of clones 2, 3 & 6 and that clones 2 & 6 produced the highest level of tTa expression (lanes 3 & 5). Clone 1 appeared to produce an uncleaved form of tTa/2A/neo^r, with clone 3 only being partially cleaved. To determine whether the tTa which was expressed from the cell lines was functional, luciferase & β -galactosidase assays were performed. Briefly, the 4 tTa expressing cell-lines were transfected either with plasmids encoding β -galactosidase or luciferase, under control of the TRP, in the absence of tetracycline. 48 hrs post-transfection the cells were assayed for luciferase activity or β -galactosidase activity as detailed in the methods section. The results of the luciferase & β -gal. assays are shown in table 5 and demonstrate that clone #6 produced the highest activity of luciferase and was thus used for further experiments (see section 2.3.8).

An analysis to determine whether the plasmid pEBV-tTa/neo has been retained as an episome, and whether EBNA-1 was being constitutively expressed has not been carried out due to time constraints.

2.3.9 / Attempt at the establishment of stable cell-lines which inducibly expressed N-terminally Pk-tagged SV5 HN, using the vectors pTEP4HNntPk (->->) or pTEP8HNntPk (->->) and a tTa expressing cell-line.

Having successfully isolated a 293 cell-line which constitutively expresses the tTa protein (293/tTa), an attempt was made to isolate cell-lines which would inducibly express the HNntPk protein encoded from an EBV-based plasmid. 5×10^6 293/tTa cells were transfected with 15 μ gs of the vector pTEP4HNntPk (->->) or pTEP8HNntPk (->->) in the presence of 3 μ gml⁻¹ tet^c. 72 hrs post-transfection, hygromycin B or histidinol was added to the medium for the selection of cells stably maintaining the HN plasmid. The cells are currently still undergoing selection.

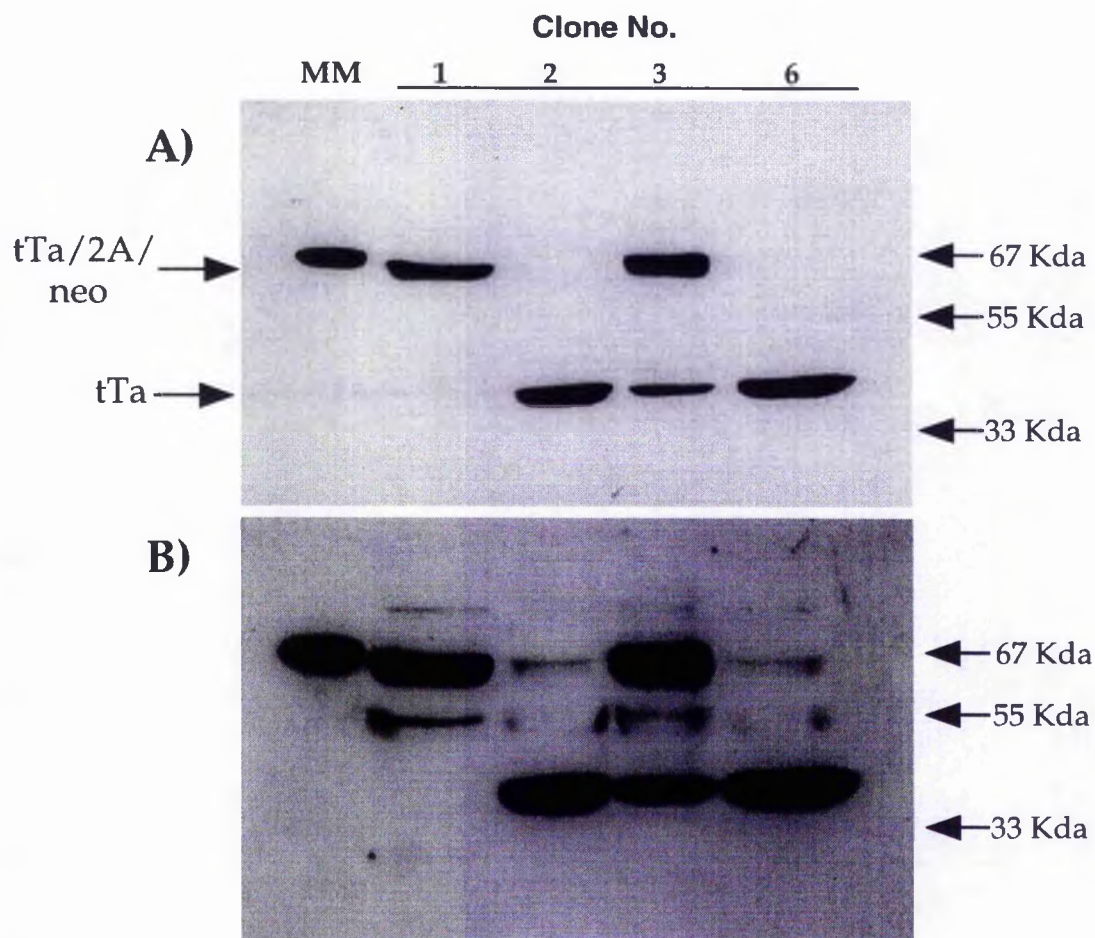


Figure 50: Western blot analysis of 4 isolated 293-based cell-lines transfected with an EBV-based plasmid producing the tTa protein.

293 cells were transfected with the vector (pEBV-tTa/neo). Drug-resistant cell-lines were isolated by selection in geneticin for two weeks (see text for details). Four clones were tested for tTa production. WCEs were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and incubated with anti-serum to the acidic domain of herpes simplex virus virion protein 16 (VP16: one half of the tTa fusion protein), followed by an incubation with anti-rabbit HRP. Detection was by ECL. (Panel A-short exposure (10 secs), panel B-long exposure (2 mins)). The positions of the bands corresponding to tTa and the uncleaved form tTa/2A/neo are indicated. MM-molecular weight markers, band detected was probably due to cross-reactivity of the anti-serum to a protein within the marker preparation.

Cell-line	Vector(s)	Luciferase Activity (RLUs/ μ g)	% cells positive for β -galactosidase
naive 293s	N/A	8	0%
naive 293s	pTRPluc.	62	N/A
naive 293s	pTRPluc & driver.	156,725	N/A
naive 293s	pCMVluc.	20,129	N/A
naive 293s	pTRP β gal.	N/A	~0.001%
naive 293s	pTRP β gal & driver.	N/A	~25%
293 EBV-tTa #1	pTRPluc or pTRP β gal.	363	~3%
293 EBV-tTa #2	pTRPluc or pTRP β gal.	454	~7%
293 EBV-tTa #3	pTRPluc or pTRP β gal.	40,746	~15%
293 EBV-tTa #6	pTRPluc or pTRP β gal.	244,824	~40%

Table 5 - Luciferase & β -galactosidase assays of 4 cell-lines which constitutively express the tTa fusion protein encoded from an EBV-based vector.

1x10⁶ 293 cells and 4 clones of tTa-expressing 293 cells were transfected with the vectors as detailed above. 48 hrs post-transfection the cells were processed for luciferase or β -galactosidase assays, as explained in the methods section. Luciferase activity was measured as relative light units per μ g of total protein (RLUs/ μ g). Plasmids pTRPluc and pTRP β gal express the luciferase and β -galactosidase genes, respectively, under control of the tetracycline responsive promoter. Plasmid pCMVluc expresses the luciferase gene under control of the hCMV major IE enhancer/promoter. The driver plasmid used was the pGEM-tTa/hygB vector. N/A-not applicable.

2.4 / Influence of the driver/responder vector-types on transient protein production from the inducible expression system.

As detailed above, a wide range of different driver/responder vectors have been constructed. Therefore, it was investigated to determine whether or not the particular type of

vector, i.e. EBV-based or non-EBV-based and (->->) or (-><-) orientated has a beneficial or detrimental effect on protein production in a transient assay.

To ensure equivalent amounts of the vectors, listed below, were transfected into the 293/E cells, 1µl of each DNA preparation (all prepared using Qiagen maxi-prep. columns) was linearised using appropriate REs. Doubling dilutions were prepared and electrophoresed through a 1% TBE agarose gel, stained using ethidium bromide and visualised by UV illumination. The fluorescent intensity was analysed using a UVP Imagestore 5000 & the SW 5000 programme (data not shown). Optical density readings of the DNA preps. at 260 & 280nm were also performed and double checked. Consistent readings of 1.78->1.81 for all preparations of DNA were obtained. Therefore, having standardised the two DNA preparations, transfections into 293/E were performed. Briefly, 5x10⁶ 293/E cells, in 75 cm² flats were transfected with the following combinations of vectors all in a 1:1 (w/w) ratio.

	EBV-based driver vector	non-EBV-based driver vector
A1	pEBV-tTa/hygB + pTEP8HNntPk (->->)	pGEM-tTa/hygB + pTEP8HNntPk (->->)
A2	pEBV-tTa/hygB + pTEP8HNntPk (-><-)	pGEM-tTa/hygB + pTEP8HNntPk (-><-)
B	pEBV-tTa/hygB + pTETHNntPk	pGEM-tTa/hygB + pTETHNntPk

- A) EBV-based responder vectors: 1- "favourable" orientation.
2- "unfavourable" orientation.
B) non-EBV-based responder vector.

Forty eight hours post-transfection, SCEs were prepared in immunoprecipitation buffer and their protein content was determined using the BCA method. An immunoprecipitation was performed, on equivalent protein amounts, using the anti-Pk mAb. The immunoprecipitate was electrophoresed through a 10% SDS-PAGE, electroblotted and probed with the anti-Pk mAb followed by an incubation with ¹²⁵Iodine-labelled anti-mouse immunoglobulin. The results were then analysed using a Fuji BAS 1000 phosphorimager (figure 52) or exposed to X-ray film (figure 51). Bands at ~ 55Kda correspond to IgH.

The results from this experiment showed that, at least in a transient system, when the responder gene was part of an EBV-based vector the nature of the driver plasmid did not

significantly influence expression levels (figure 51, groups A & B and figure 52, columns 1 & 3 and 2 & 4). However, when the responder gene was part of a non-EBV-based vector then an EBV-based driver vector appeared to produce a greater amount of protein (figure 51 lanes 5 & 6 and figure 52, columns 5 & 6).

The EBV-based responder plasmid which had the *goi* transcription unit in the "favourable" orientation (figure 51, lanes 1 & 3 and figure 52, columns 1 & 3) produced a consistently greater amount of gene product than the "unfavourable" orientation (figure 51, lanes 2 & 4 and figure 52, columns 2 & 4) - approximately 2 fold - when either type of driver plasmid was co-transfected. Furthermore, both orientations of EBV-based responder vectors (figure 51, group A & B and figure 52, columns 1,2,3 & 4) produced greater amounts of gene-product than the non-EBV-based responder plasmid - approximately 3-8 fold (figure 51, group C and figure 52, columns 5 & 6).

2.5 / Summary.

- a) No expression of C-terminally Pk-tagged HN could be detected from EBV-based vectors, when transcription was driven by the hCMV enhancer/promoter or RSV LTR, in either transiently transfected cells or from drug-resistant cell-lines.
- b) Production of Pk-tagged HN could be detected from the pcDNA-based vectors when transcription was driven by the T7 ϕ 10 promoter, but not from the hCMV promoter.
- c) Production of N-terminally Pk-tagged HN, but not C-terminally Pk-tagged HN, could be detected when expression was driven from the tetracycline-responsive promoter.
- d) In a transient expression analysis, the use of EBV-based driver & responder plasmids resulted in significant increases in Pk-tagged HN production in 293/E cells.
- e) Transfection of EBV-based plasmids into an EBNA-1 expressing 293 cell-line, followed by subsequent drug-selection, did not increase the number of drug-resistant colonies produced when compared to a similar transfection/selection procedure performed in native 293 cells.
- f) EBV-based and non-EBV-based plasmids encoding the genes for tTa and hyg B PT, separated by the FMDV 2A sequence, were constructed and shown to co-express functional tTa and hyg B PT.

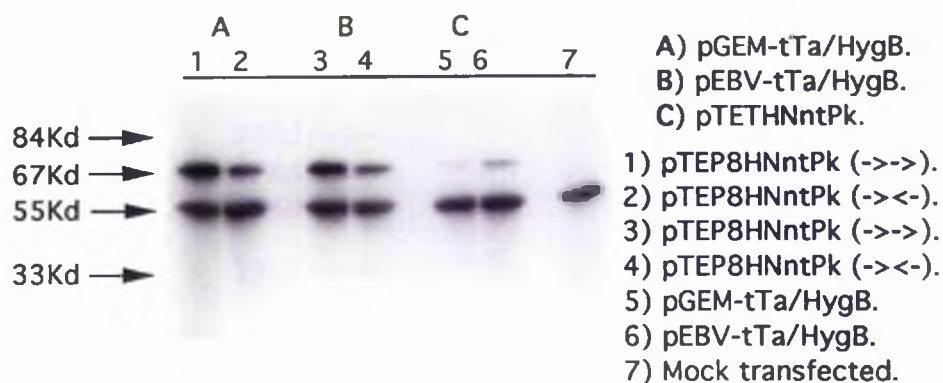
g) A tTa expressing cell-line was isolated which expressed high-levels of functional tTa protein from an EBV-based vector. Episomal maintenance of the plasmid has not yet been investigated.

h) A cell-line which inducibly expresses Pk-tagged HN has not yet been produced.

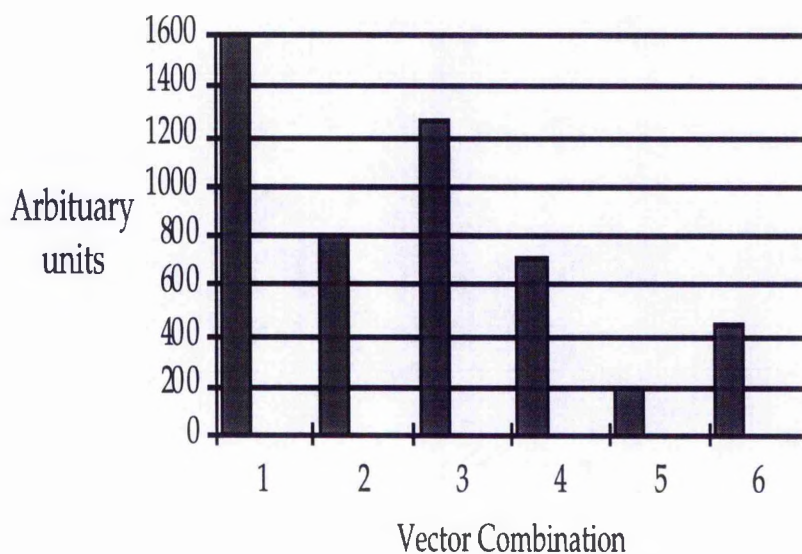
Figure 51: Phosphoimaging analysis of the influence of vector-type on SV5 HNntPk gene expression in a transient assay.

293 /E cells were transfected with the shown vector combinations in a 1:1 ratio, in the absence of tetracycline. 48 hours post-transfection, SCEs were prepared in immunoprecipitation buffer, and the protein content was standardised using the BCA assay. An immunoprecipitation was performed, on equivalent protein amounts, using the anti-Pk mAb. The immunoprecipitate was electrophoresed through a 10% SDS-PAGE, electroblotted and probed with the anti-Pk mAb followed by an incubation with ¹²⁵Iodine-labelled anti-mouse immunoglobulin and exposure to X-ray film (panel A) The results were then analysed using a Fuji BAS 1000 phosphorimager (panel B). Results are given as arbitrary units.

A)



B)



Vector Combination	Numerical Value
EBV-responder (->->) / GEM-tTa driver	1
EBV-responder (-><-) / GEM-tTa driver	2
EBV-responder (->->) / EBV-tTa driver	3
EBV-responder (-><-) / EBV-tTa driver	4
Non-EBV-responder / GEM-tTa driver	5
Non-EBV-responder / EBV-tTa driver	6

SECTION 3 : A NOVEL APPROACH FOR THE AMPLIFICATION OF THE COPY NUMBER OF EPISOMES STABLY MAINTAINED WITHIN PRIMATE CELL-LINES.

EBV-based episomal vectors are maintained at a low copy number (between 1- 50 copies per cell) in the primate cell nucleus. Transient expression systems utilising the Simian Virus 40 (SV40) Large T Antigen (LTAg), as expressed constitutively in Cos cells, and the SV40 origin of replication (*ori*), as located on the vector DNA, can produce copious amounts of product due to the rapid replication of the vector DNA by the LTAg, resulting in up to 10,000 copies per cell. Replication continues unchecked until cell death occurs approximately 3-4 days post-transfection. It was envisaged that a system could be developed whereby stable cell-lines would be produced, which inducibly express the SV40 LTAg under tTa inducible control. Such stable cell-lines, when in a repressed state, could then be used for the isolation of cells which constitutively express a gene of interest encoded from a stably maintained SV40 *ori*-containing EBV-based plasmid. Upon induction of the LTAg, the LTAg would amplify the copy number of the SV40 *ori*-containing EBV-based plasmids with the expectation that , upon amplification, significant increases in the level of recombinant protein production may be achieved. As a model to investigate the feasibility of such an amplification system, a non-toxic protein - the SV5 P protein - which could thus be expressed constitutively from the hCMV major IE enhancer/promoter, was employed.

3.1 / Construction of an SV40 *ori*-containing EBV-based vector which constitutively expressed the SV5 P protein (figure 53).

The vectors pCEP8/P (->->) & (-><-) were constructed as follows. The transcription unit encompassing the hCMV IE/HNPk/pA was excised from the plasmid pCEP4HNctPk (->->) using the RE Sal1 and ligated into the similarly digested vector pREP8 to produce the plasmid pCEP8HNctPk (->->). The gene encoding the SV5 P protein was excised from the plasmid pGEM/P using the REs Kpn1/Sal1 and ligated into the Kpn1/Xho1 digested pCEP8HNctPk (->->) vector to produce the desired plasmid pCEP8/P (->->). The vector was tested for P protein expression as detailed below. pCEP8/P (->->) was modified by a

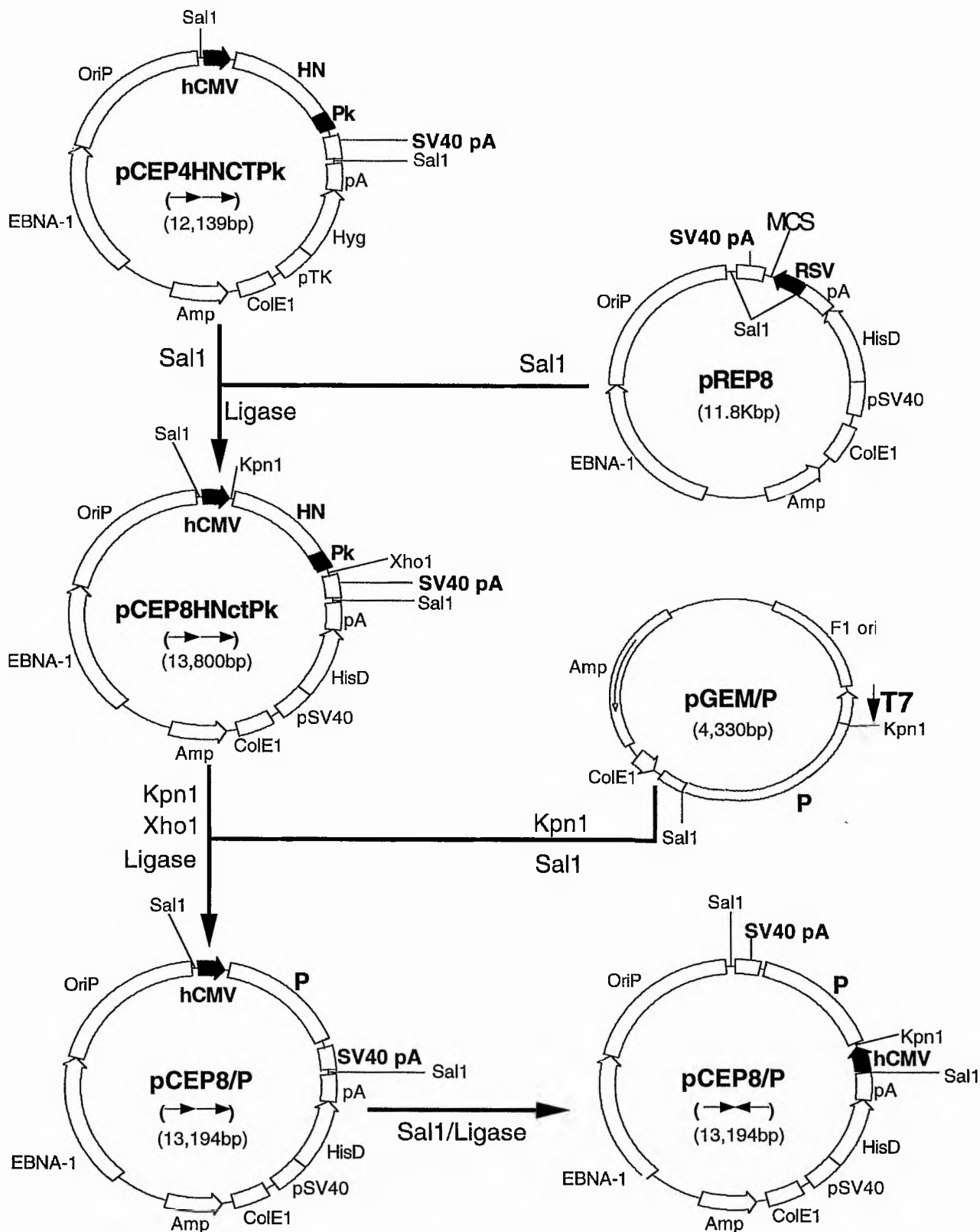


Figure 53: Construction of LTag-amplifiable EBV-based vectors which constitutively express the SV5 P protein.

The plasmid pCEP4HNctPk (->->) was digested with the RE Sal1 and the transcription unit containing the hCMV/HN/pA was ligated into the similarly digested vector pREP8. This new plasmid, pCEP8HNctPk (->->) was digested with the REs Kpn1 & Xho1 into which the gene encoding the SV5 P protein was ligated, thus producing the vector pCEP8/P (->->). This plasmid was further modified by a Sal1 digestion, then ligation, to produce the vector pCEP8/P (-><-).

digestion with the RE Sal1 and subsequent ligation of the two fragments. Constructs in which the hCMV IE/P/pA transcription unit was ligated in the opposite orientation to the (->->) parental vector were isolated, and named pCEP8/P (-><-).

3.2 / Transient protein production from the vector pCEP8/P (->->).

1x10⁶ 293/E cells were transfected with the vectors pCEP8/P (->->), pGEM-tTa/-hygB & pTET/V (expresses the SV5 V protein which is also recognised by the anti-Pk mAb and is used as a positive control for the Western blot) and pCEP4 (negative control - mock transfected). 48hrs post-transfection, WCEs were prepared in x1 disruption buffer and separated through a 10% SDS-PAGE, electroblotted then probed with the anti-Pk mAb. Figure ³~~54~~ shows that the 46Kda P protein was clearly detected from WCEs prepared from pCEP8/P (->->) transfected cells (lane 2), along with the 23Kda V protein (lane 1). The cytoplasmic localisation of the P protein was demonstrated by indirect immunofluorescence analysis (using the anti-Pk mAb) on transfected 293/E cells, the result of which is depicted in figure ⁴~~55~~.

3.3 / Manipulations to increase the expression of the SV5 P protein encoded from EBV-based vectors.

3.3.1 / Influence of the orientation of the SV5 P gene transcription unit on the production of P protein in a transient expression analysis.

To ensure equivalent amounts of the vectors pCEP8/P (->->) & (-><-) were transfected into the 293/E cells, 1µl of each DNA preparation (both prepared using Qiagen maxi-prep. columns) was linearised using the RE Sph1. Doubling dilutions were prepared and electrophoresed through a 1% TBE agarose gel, stained using ethidium bromide and visualised by UV illumination. The fluorescent intensity was analysed using a UVP Imagestore 5000 & the SW 5000 programme (Data not shown). Optical density readings of the DNA preps. at 260 & 280nm were also performed and double checked. Consistent readings of 1.78->1.80 for both preparations of DNA were obtained.

Therefore, having standardised the two DNA preparations, transfections into 293/E cells, grown to 40% confluency in 35mm petri dishes, were performed. 48hrs post-

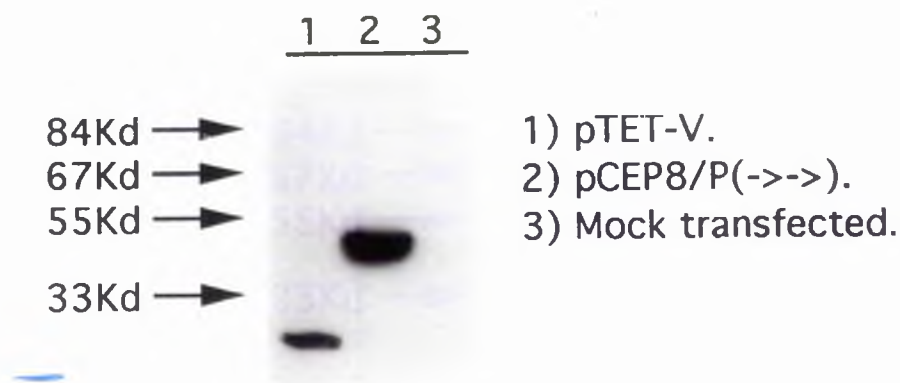


Figure 53: Transient expression of the SV5 P protein from an SV40ori -containing EBV-based vector.

293/E cells were transfected with the plasmids pTET/V & pGEM-tTa/hygB (lane 1) and pCEP8/P (->->) (lane 2) or were mock transfected (lane 3). 48hrs post-transfection WCEs were prepared in x1 disruption buffer and then separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by detection using ECL. Included was a WCE from cells transiently expressing the SV5 V protein, which is also recognised by the anti-Pk mAb and thus acted as a positive control (lane 1).

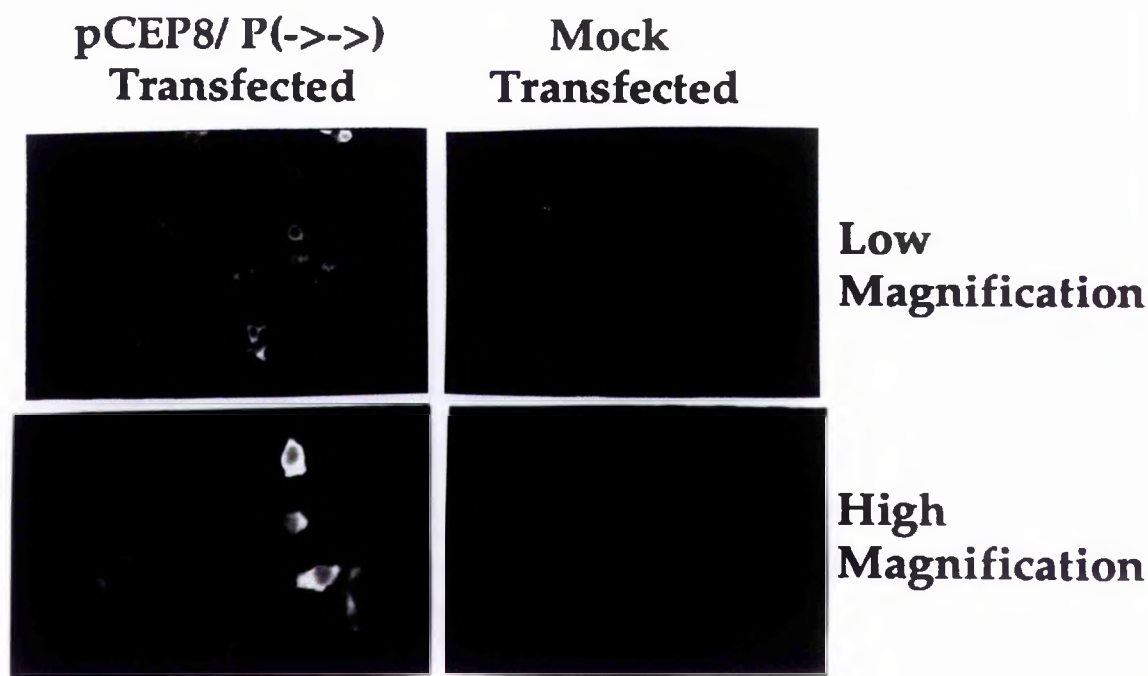


Figure 54: Immunofluoresence analysis of pCEP8/P transfected 293/E cells.

293/E cells were transfected with the vector pCEP8/P (->->) or mock transfected. 48hrs post-transfection the cells were fixed & permeabilised then incubated with the anti-Pk mAb followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant. Low magnification x10, high magnification x40.

transfection WCEs were prepared, in 100µls of 1% NP40 in PBS, by sonication and the samples standardised for protein content by the BCA assay. The standardised samples were then adjusted to x1 disruption buffer and 5µl, 10µl & 15µl of equivalent protein amounts were separated through a 10% SDS-PAGE, electroblotted, probed with the anti-Pk mAb, followed by secondary detection with ¹²⁵Iodine-labelled anti-mouse immunoglobulin. The blots were then analysed using a Fuji BAS 1000 phosphorimager and exposed to X-ray film (figure 5⁵~~6~~). The results show that for the SV5 P protein, produced in a transient expression assay, the vector pCEP8/P (->->) produced more P protein compared to the vector pCEP8/P (-><-). Analysis by phosphorimaging quantitated this difference to be ~3 fold (data not shown).

3.3.2 / Effect of synchronising the mammalian cells at the G2/M phase of the cell-cycle prior to transfection.

Experiments carried out by Teshigawara & Katsura (1992) demonstrated that synchronisation of the EBV transformed B cell-line, Raji, at the G2/M phase of the cell-cycle immediately prior to transfection, significantly increased the level of IL-2 receptor β-chain production, encoded on an EBV-based plasmid, by approximately 10-fold. Therefore, 293/E cells grown in 35mm petri dishes to 40% confluency, were synchronised at the G2/M phase of the cell-cycle, as described in materials & methods. The synchronised cells, and an unsynchronised population grown to a similar density, were transfected with the vector pCEP8/P (->->). 48hrs post-transfection, WCEs were prepared in 100µls of 1% NP40 in PBS by sonication and the samples standardised for protein content by the BCA assay. The standardised samples were then adjusted to x1 disruption buffer and equivalent protein amounts were separated through a 10% SDS-PAGE, electroblotted, probed with the anti-Pk mAb, followed by secondary detection with ¹²⁵Iodine-labelled anti-mouse immunoglobulin. The blots were then analysed using a Fuji BAS 1000 phosphorimager and exposed to X-ray film (figure 5⁶~~7~~). No substantial difference between P protein detection levels were observed between the synchronised and unsynchronised cell populations, in a transient expression analysis.

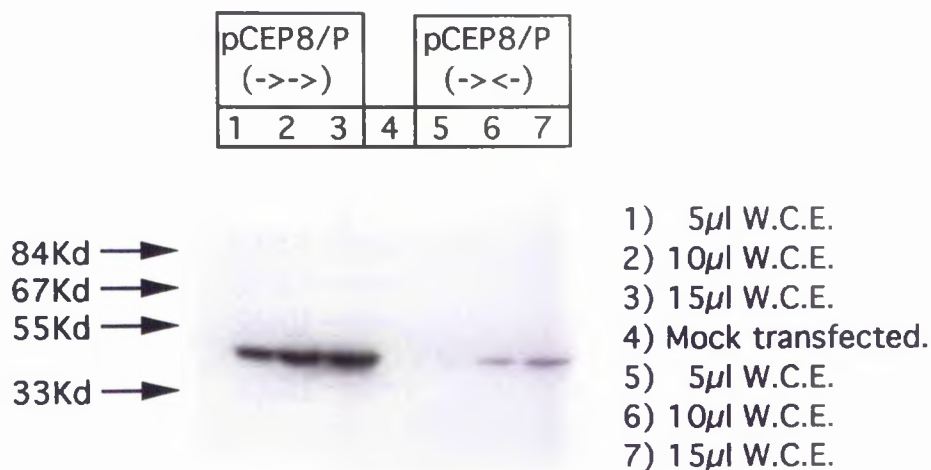


Figure 55: Effect of transcription unit orientation upon the transient expression level of the SV5 P protein from an SV40*ori* -containing EBV-based vector.

293/E cells were transfected with the plasmids pCEP8/P (->->) (lanes 1-3) and pCEP8/P (-><-) (lanes 5-7) or were mock transfected (lane 4). 48hrs post-transfection, WCEs were prepared in immunoprecipitation buffer and their protein content was standardised using the BCA assay. Equivalent amounts of protein were then separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by 125-Iodinated anti-mouse Ig and exposure to X-ray film.

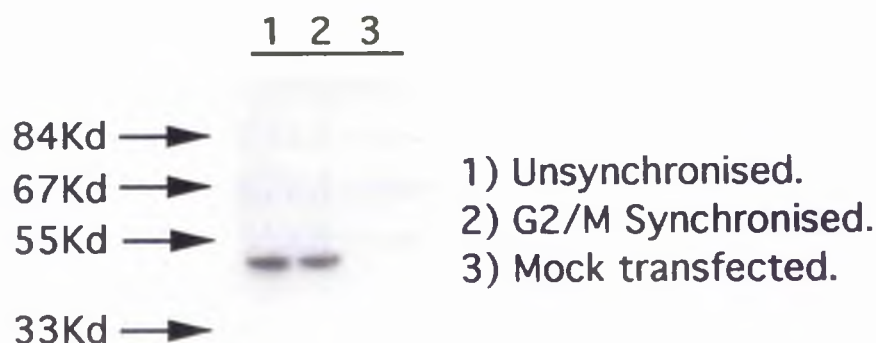


Figure 56: Effect of cell synchronisation upon the transient expression level of the SV5 P protein from an SV40ori -containing EBV-based vector.

293/E cells were synchronised at the G2/M phase of the cell-cycle by an overnight treatment with thymidine followed by an overnight incubation with Hoechst 33342. Synchronised and unsynchronised cells (grown to a similar density) were transfected with the plasmid pCEP8/P (->->) (lanes 1 & 2) or were mock transfected (lane 3). 48hrs post-transfection WCEs were prepared in immunoprecipitation buffer and their protein content was standardised using the BCA assay. Equivalent amounts of protein were then separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-Pk mAb followed by 125-Iodinated anti-mouse Ig and exposure to X-ray film.

3.4 / Construction of the pCT7-based vector which encodes the SV40 LTA_g.

Having successfully constructed appropriate SV40-*ori*-containing EBV-based vectors which were shown to constitutively express the SV5 P protein, or inducibly express the N-terminally Pk-tagged SV5 HN glycoprotein, the plasmid which inducible expresses the SV40 LTA_g was constructed. However, prior to producing a vector which would inducibly express the SV40 LTA_g, a vector which expresses the LTA_g under the hCMV IE enhancer/promoter or the T7 ϕ 10 promoter was constructed. The ORF for the SV40 LTA_g was PCR amplified from the vector pSVori using primers containing the REs EcoR1 or Sph1. The PCR product was digested with the aforementioned REs and ligated into the similarly digested plasmid pCT7, to produce the vector pCT7LTA_g, (figure 58).

3.5 / Transient production of the SV40 LTA_g from the vector pCT7LTA_g.

Expression from the vector pCT7LTA_g was initially detected by an indirect immunofluorescence analysis. Briefly, 293/E cells were grown on glass coverslips and then transfected with the plasmids pCT7LTA_g or pcDNA1/amp (pCT7-mock transfected). 48hrs post-transfection the cells were processed for immunofluorescence using the anti-LTA_g mAb. Figure 59 clearly shows the nuclear localisation of LTA_g in cells transfected with the plasmid pCT7LTA_g.

To confirm that the LTA_g was of the correct molecular weight, a Western blot analysis of WCEs was performed. Briefly, 293/E cells grown in 35mm petri dishes were transfected with the plasmids; pCT7LTA_g, pSVori and pCT7 (mock transfection) as detailed in the materials & methods section. At 48hrs post-transfection, WCEs were prepared in x1 disruption buffer, including a WCE from Cos-7 cells (positive control). Figure 60 (panel A) shows the result of separating the WCEs through a 10% SDS-PAGE, electroblotting and probing with the anti-LTA_g mAb. A band of ~90Kda is clearly detectable from extracts of Cos-7 cells and from cells transfected with the pSVori vector, albeit at a much lower level. However, extracts from cells transfected with the vector pCT7LTA_g showed only a very weak signal for an ~90Kda band with a prominent signal at 33Kda and minor signals at ~55Kda & 40Kda. All of these latter signals were not present in the Cos-7 cell extract or the

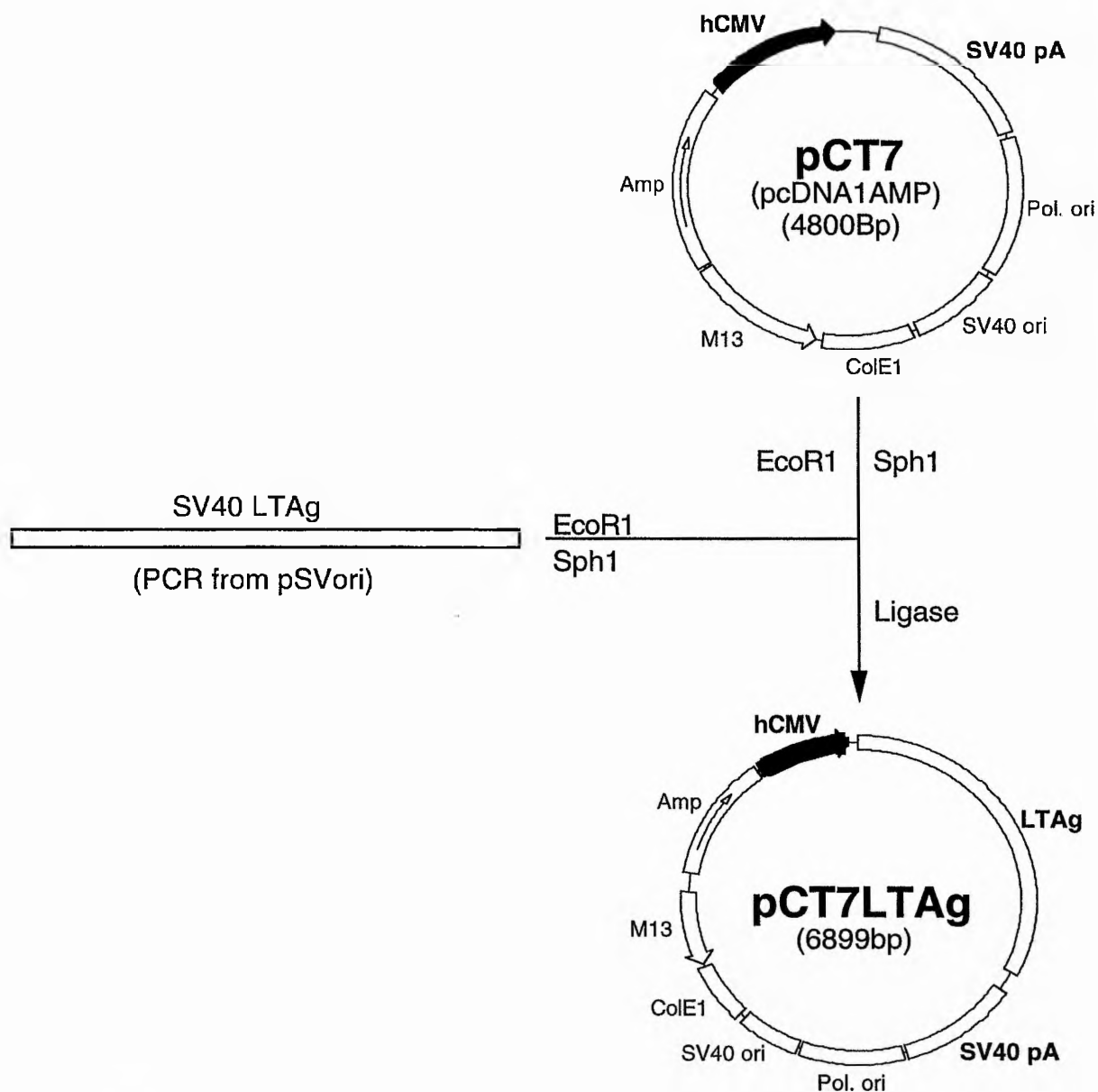


Figure 58: Construction of the vector for the transient expression of the SV40 LTA_g.

The DNA sequence encoding the open reading frame of the LTA_g was PCR amplified from the vector pSVori, employing PCR primers which contain either EcoR1 or Sph1 RE sites. The PCR product was digested with the REs Sph1 & EcoR1, and ligated into the similarly digested plasmid pCT7 to produce the vector pCT7LTA_g, from which transient expression of the LTA_g can be driven by the hCMV promoter.

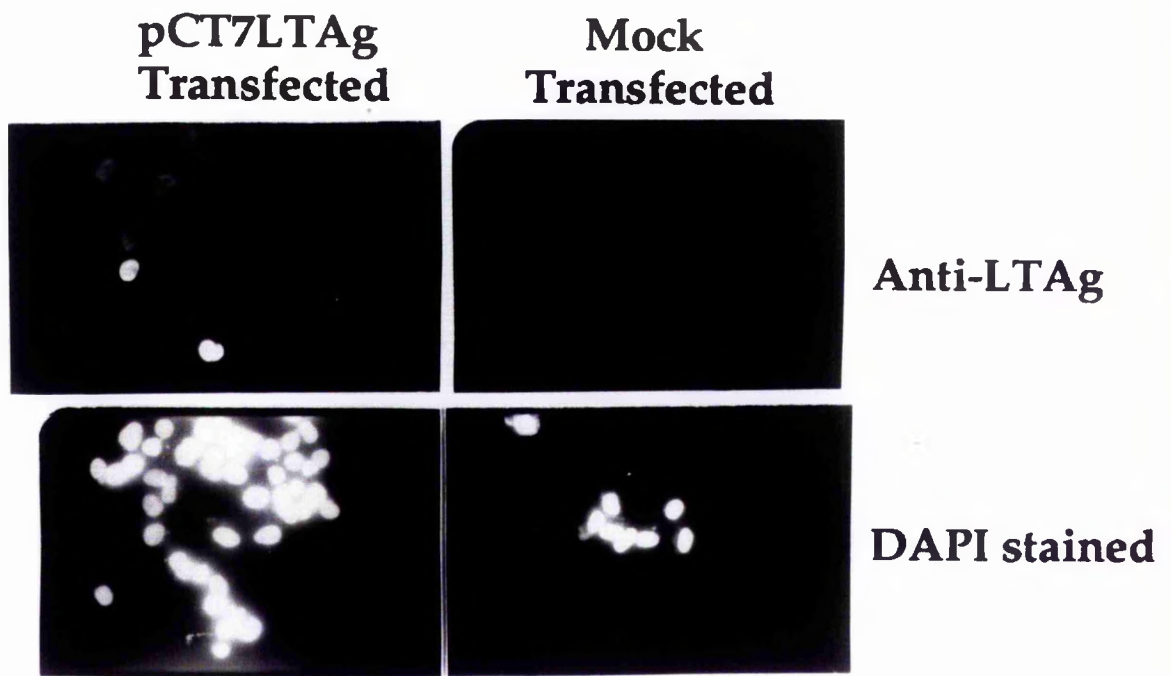
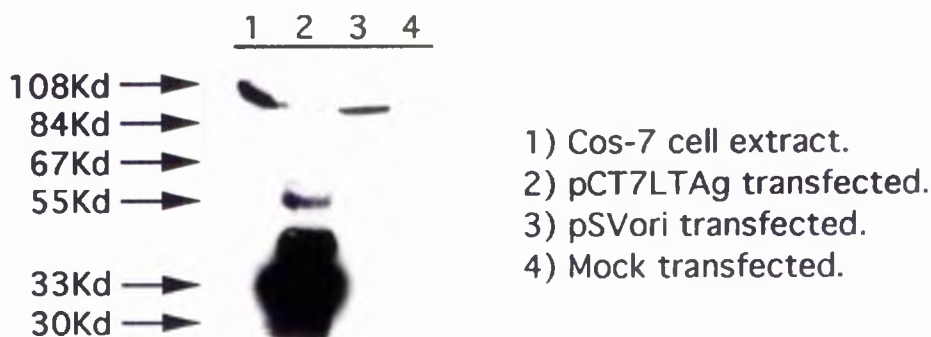


Figure 58: Immunofluoresence analysis of pCT7LTA transfected 293/E cells.

293/E cells were transfected with the vectors pCT7LTA or pcDNA1/Amp (mock transfected). 48hrs post-transfection, cells were fixed & permeabilised then incubated with the anti-LTA mAb (PAb 416) followed by an incubation with Texas Red conjugated anti-mouse Ig. Exposure times are constant. The cell nuclei were stained with DAPI.

(A)



(B)

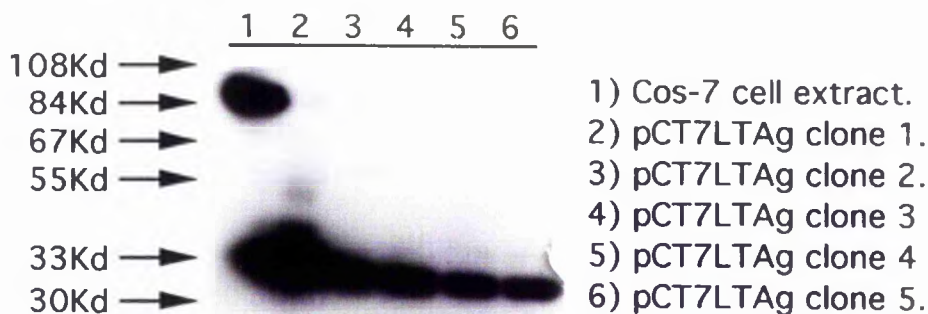


Figure 59: Transient Expression of SV40LTA from hCMV IE promoter in 293/E cells

(A) 40% confluent 35mm petris of cells were transiently transfected with 3 μ g of either; pCT7LTA (lane 2), pSVori (+ve for LTA, lane 3) or pcDNA1/amp (mock transfected. -ve for LTA, lane 4). 48hrs post transfection, WCEs were prepared and separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-SV40 LTA mAb - PAb 416. Detection was by ECL. A sample of WCE derived from Cos-7 cells was also incorporated as a positive control (lane 1). **(B)** A 40% confluent 35mm petri of cells was transiently transfected with 3 μ g of DNA (pCT7LTA, lanes 2-6) isolated from 5 different transformants showing positive restriction endonuclease digestion pattern. 48hrs post transfection, WCEs were prepared and analysed as above.

extract from the pSVori transfected cells. However, since the ORF was PCR amplified from the vector pSVori, mutations could have been introduced into the LTA_g ORF sequence resulting in the predominance of the lower molecular weight bands. Therefore, DNA minipreps from 5 separate transformants (which gave a positive diagnostic RE digestion) from the same ligation reaction were prepared. A transient expression experiment in 293/E cells was performed using the DNA minipreps, with the transfected cells being processed as detailed above. Figure 3.7 (panel B) shows that for all the clones tested, the ~33Kda, 40Kda and 55Kda bands were present, with only a weak signal corresponding to the full length LTA_g being detected (note- a longer exposure was needed to observe the 55Kda and 90Kda bands detected from the WCE of cells transfected with clones 2-5. Data not shown).

3.6 / Establishment of cell-lines which constitutively expressed the SV5 P protein encoded from an EBV-based vector.

Although the inducible LTA_g expressing construct has, as yet, not been constructed, the ability to increase protein production after amplification of a stably maintained low copy number episome can be tested by the transient transfection of LTA_g into a stable cell-line possessing an episome. For this purpose, cell-lines which constitutively express the SV5 P protein encoded on the vector pCEP8/P (->->) were produced by transfection of naive low passage no. 293 cells followed by histidinol selection, in a procedure similar to that described for the isolation of tTa expressing cell-lines. Initially five clones (#2, #3, #4, #5 & #11) with different properties were isolated from the initial screening of 24 clones by immunofluorescence analysis using the anti-Pk mAb (data not shown). Clone #2 expresses P protein in a high proportion of cells with low expression (HNLE). Clones #3 & #4 express the P protein in a high proportion of cells with high expression (HNHE). Clone #5 expresses P protein in a high proportion of cells with intermediate expression (HNIE). Clone #11 expresses P protein in a low proportion of cells with high expression (LNHE).

3.7 / Transient expression of SV40 LTA_g in cell-lines which express the SV5 P protein encoded from an SV40 ori -containing EBV-based vector.

All five P-expressing cell-lines were used in a transient transfection assay to determine whether LTA_g expression could increase P protein expression. The five cell-lines

were grown, in duplicate, to ~40% confluency in 35mm petri dishes where-upon each cell-line was transfected with the plasmid pSVori or mock transfected. 72hrs post-transfection the cells were harvested by washing twice with PBS followed by lysis with immunoprecipitation buffer. Samples were standardised for protein content using the BCA assay and made up to x1 disruption buffer. Equivalent protein amounts of WCE were separated through a 12% SDS-PAGE, electroblotted and probed with the anti-Pk mAb and anti-LTA_g mAb, followed by detection using ECL. Figure 6⁹ shows that cell-lines #2 (lanes 1 & 2) & #3 (lanes 3 & 4) may show a very small increase in P protein expression (46Kda band) after transient expression of the SV40 LTA_g (~90Kda band), however, cell-line #11 (lanes 9 & 10) may be showing a small decrease in P protein expression. Therefore, overall, no substantially large increase in P protein expression was observed following transfection of P-expressing cell-lines with the plasmid pSVori.

3.8/ Summary.

- 1) EBV-based vectors which possessed an SV40 *ori* were constructed, and shown to produce the SV5 P protein when expression was driven by an hCMV IE enhancer/promoter. This expression, in a transient expression analysis, was shown to be maximal when the orientation of the P gene transcription unit was similar to that of the EBNA-1 gene transcription unit.
- 2) Synchronisation of 293 cells prior to transfection did not result in a detectable increase in protein production.
- 3) An attempt at the construction of a plasmid which expressed the SV40 LTA_g under control of an hCMV IE enhancer/promoter, resulted in the production of a nuclear-localised protein, but which primarily possessed a mwt. of only ~33 Kda. A small amount of protein was detected at the expected mwt. of ~90 Kda. The construction of a plasmid which inducibly expressed the SV40 LTA_g was not carried out.
- 4) Transfection of 5 separate SV5 P producing cell-lines (in which the P gene was encoded on an SV40 *ori*-containing EBV-based plasmid) with the LTA_g expressing vector pSV *ori*, did not result in a detectable increase in P protein expression.

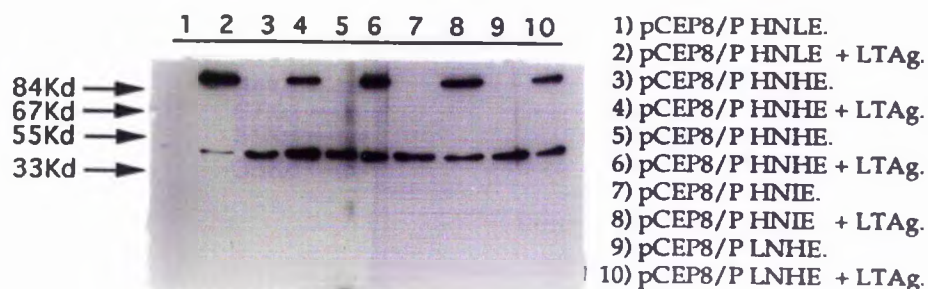


Figure 60: Transient expression of the SV40 LTag in cell-lines which constitutively express the SV5 P protein from an SV40*ori* -containing EBV-based episome.

Five 293 cell-lines which constitutively express the SV5 P protein (see text for details) were transfected with the plasmid pSVori (lanes 2, 4, 6, 8 & 10) or were mock transfected (lanes 1, 3, 5, 7 & 9). 72hrs post-transfection WCEs were prepared in immunoprecipitation buffer and their protein content was standardised using the BCA assay. Equivalent amounts of WCE were separated through a 10% SDS-PAGE under reducing conditions, electroblotted and probed with the anti-LTag mAb and anti-Pk mAb followed, by detection using ECL.

DISCUSSION

1/ Pk-tag as an epitope-tag for the detection and purification of recombinant proteins.

1.1 Analysis of transiently expressed Pk-tagged HN & F.

The mammalian expression system chosen for the analysis of Pk-tagged HN & F processing was the vaccinia virus - T7 RDRP system. This system was chosen due to the ease of use, the high level of protein production which occurs, and the brief time that is required to obtain protein production. Vectors were constructed, termed pT7ntPk & pT7ctPk, which contained oligonucleotides coding for the Pk-tag which is subsequently positioned at the N or C-terminus, respectively, of an inserted gene. The SV5 HN & F genes or their truncated forms (Δ), were PCR-amplified from appropriate plasmids and inserted into the above vectors, to produce constructs which expressed the following proteins from the T7 ϕ 10 promoter; HNntPk, HNctPk, FctPk, Δ HNntPk, Δ HNctPk & Δ FctPk.

1.1.1 / Analysis of full length Pk-tagged HN & F.

Western blot analysis of WCEs from transfected vTF7-3 infected 293 cells (figure 21) demonstrated that products of the expected mwt. were detected, using the anti-Pk mAb, and therefore, the addition of the Pk-tag did not prevent the expression of the Pk-tagged glycoproteins in mammalian cells. Whether the presence of the Pk-tag had significantly reduced the expression of Pk-tagged glycoproteins when compared to the native glycoproteins has not been extensively investigated due to the fragility of the vaccinia virus

infected 293 cells (whereby repeated addition/aspiration of medium resulted in the substantial disintegration of the cell monolayer), thereby rendering radio-labelling experiments redundant. However, a comparison between the intensity of the immunofluorescence, when incubated with anti-HN mAbs, observed between Pk-tagged HN (figures 24 & 26) and native HN (figure 23) would indicate that there is not a substantial reduction in expression of the Pk-tagged form of HN (a western blot analysis could not be performed due to the lack of a mAb which recognised Western blotted SV5 HN).

The Pk-tag epitope does not appear to be cleaved from the Pk-tagged proteins by cellular or vaccinia virus proteases as products could be detected when probed with the anti-Pk mAb (figure 21). Furthermore, immunofluorescence analyses shown in figures 24 & 26 demonstrate a comparable level of detection of Pk-tagged HN using either the anti-HN or anti-Pk mAbs, which would tend to argue against there only being a small subset of proteins which possess an intact Pk-tag. However, a definitive analysis has not been carried out to determine whether or not a small proportion of the expressed proteins have lost the Pk-tag epitope.

The native SV5 F glycoprotein is a type I glycoprotein and possesses a cleavable signal sequence. The successful expression of the Pk-tagged form of F to produce a protein of the expected mwt., would infer that the presence of the Pk-tag at the C-terminus of the molecule did not prevent cleavage of the signal sequence by signal peptidase. Moreover, the addition of the Pk-tag to the C-terminus of F did not prevent the post-translational cleavage of the precursor protein F₀ to the disulphide-linked F₁ & F₂ subunits.

Native HN & F undergo N-linked glycosylation which adds approximately 6-10 Kda to the apparent mwt. of the proteins. Pk-tagged HN and F were shown, using PNGase F digestion analysis (figure 22) or tunicamycin treatment, to be post-translationally modified by the addition of 6-10 Kda of N-linked oligosaccharides. Furthermore, native HN & F are expressed at the cell-surface of infected cells. Successful transport of Pk-tagged HN & F to the surface of transfected cells was demonstrated using immunofluorescence analysis (figures 24, 26 & 27). However, experiments performed by Parks & Lamb (1990) on the SV5 HN protein demonstrated that the alteration/substitution of the N-terminal cytoplasmic tail may alter

the transport of the mutant molecule through the exocytotic pathway resulting in a reduced transport rate or cessation of transport to the cell-surface. Although Pk-tagged HN & F were clearly transported to the cell-surface and were orientated in the expected manner, with HNctPk possessing the Pk-tag epitope on the exterior of the cell and HNntPk & FctPk possessing the Pk-tag on the interior of the cell, their rate of transport has not been determined. Since native HN is internalised from the cell-surface it may be prudent to investigate whether or not the Pk-tagged forms are also internalised. In terms of maximising recombinant protein production, a possible inhibition of SV5 HN internalisation, by the presence of the Pk-tag, may result in a greater yield of the protein due to retention at the cell-surface.

Native monomeric HN formed dimers which associated to produce a mixture of covalently and non-covalently linked tetramers. Preliminary evidence which suggested that Pk-tagged HN formed an oligomeric structure was demonstrated by employing non-reducing SDS-PAGE (figure 28). This analysis demonstrated that the Pk-tagged forms of HN produced mwt. species which have a similar mwt. to dimeric (~ 140 Kda) and tetrameric (~289 Kda) forms of native HN (Ng *et al.*, 1989). However, as above, the rate of oligomeric formation was not determined. Furthermore, perhaps sucrose density centrifugation analysis should be performed to further support the evidence for the oligomeric structure for Pk-tagged HN.

Using chemical cross-linking analysis, Russell *et al.* (1994) demonstrated that the SV5 F protein probably formed non-covalently linked trimers. Using a similar method, it was demonstrated that, in the majority of the detected molecules, FctPk formed dimers and possibly trimers (figure 29). Smearing of the higher mwt. bands could not be alleviated by altering gel composition and electrophoresis times, and therefore may be a reflection that the proteins were not immunoprecipitated after chemical cross-linking, thereby possibly allowing the detection of promiscuously cross-linked species.

One of the most important features when the Pk-tag was added to the termini of HN & F, in terms of vaccine development & structural studies, was to determine if the presence of the Pk-tag altered the conformation of HN & F such that they could no longer be

recognised by conformationally-sensitive (cs) mAbs, which were raised against the native proteins. Immunoprecipitation analysis (figure 30) demonstrated that Pk-tagged HN was indeed recognised by all of the cs anti-HN mAbs employed in the experiment. Furthermore, mAb 4b has been shown (Ng *et al.*, 1989) to bind only to the dimeric form of HN, thus supporting the evidence for an oligomeric structure for Pk-tagged HN. Immunofluorescence analysis (figure 27) demonstrated that the addition of the Pk-tag to the C-terminus of the F protein did not result in a loss of binding to the cs anti-F mAb.

1.1.2 / Analysis of truncated Pk-tagged HN & F.

Parks & Lamb (1990) have demonstrated that a truncated form of the SV5 HN protein (HN-F) could be secreted from CV-1 cells which were infected with an SV40 recombinant virus. The HN protein was converted into a secretable form by replacing the N-terminus signal/anchor domain with the fusion related ectodomain from the SV5 F protein. Approximately 40% of HN-F was secreted from the cells and found in a dimeric form (with a monomeric relative mwt. of ~66Kda), which retained antigenicity to cs anti-HN mAbs. The remainder of the HN-F was found to consist of dimers and monomers which were not recognised by cs anti-HN mAbs. The successful secretion, from mammalian cells, of truncated HN protein derived from PIV-3 has also been demonstrated by Lehman *et al.* (1993).

Therefore, it was determined whether or not Pk-tagged forms of truncated HN could also be secreted from mammalian cells. For this purpose, the vectors pT7 Δ HNctPk & pT7 Δ HNntPk were constructed which possessed the putative cleavable signal sequence from the EBV gp220/360 glycoprotein. Western blot analysis of WCEs from transfected vTF3-7 infected 293 cells (figure 33) demonstrated that the majority of the detected Δ HNntPk & Δ HNctPk molecules were found as a doublet band of ~60-62 Kda, which were not observed by Parks & Lamb (1990) during expression of the HN-F molecule. However, a minor band at ~68Kda was detected from WCEs of cells transfected with the pT7 Δ HNntPk vector which approximates to the expected mwt. for glycosylated monomeric Δ HNntPk. Indeed N-linked glycosylation of this minor ~68 Kda species was confirmed by PNGase F digestion analysis

(figure 36). However, no corresponding reduction in molecular weight was observed for the major doublet species. The minor ~68 Kda species was also shown to bind to a pool of cs anti-HN mAbs as determined by an immunoprecipitation analysis, but was not detected in the medium (figure 34). Moreover, this minor species was absent from WCEs of cells transfected with the pT7 Δ HNctPk vector, for reasons that are unclear, but could be related to the nature of the amino acids present immediately after the signal sequence, which differs between the two Δ HN molecules (see figure 62).

A lack of detectable secretion of Δ HNntPk & Δ HNctPk could be due to various factors, including; cleavage of the Pk-tag resulting in non-detection, lack of (or inefficient) cleavage of the putative signal sequence and, incomplete formation of transport competent molecules. Due to the inability to perform pulse-chase experiments on vTF7-3 infected 293 cells (as mentioned earlier), it could not be determined whether non-secretion was due to a deficiency in detection when probed with the anti-Pk mAb.

A deficiency (or inefficiency) in cleavage of the EBV-derived signal sequence could lead to a lack of detectable secretion into the medium. Indeed, it has recently been elucidated that the putative signal sequence used in this experiment may not be efficiently cleaved and that a modified form has been produced which results in a more efficient cleavage (M. Mackett, personal communication). This modified form of the EBV gp220/360 signal sequence, or other cleavable signal sequences such as those derived from the influenza virus HA protein (Madison & Bird, 1992) or tissue plasminogen activator protein (Rhodes *et al.*, 1994), could be utilised in the construction of future vectors encoding secretory forms of Pk-tagged HN & F. However, if the hypothesis that the lack of secretion was due to non-existent or poor cleavage of the signal sequence, then full length Pk-tagged Δ HN molecules should be produced and would perhaps be expected to behave similarly to the Pk-tagged HN molecules described earlier. Clearly this was not observed.

Parks & Lamb (1990) demonstrated that the formation of the HN-F dimer was a prerequisite for transport to the *medial* golgi, but that the formation of complex oligosaccharide chains in the *medial* golgi is not a prerequisite for secretion. In their experiments, Parks & lamb used a different signal sequence for the secretion of the truncated

A)

GAATTC ATG GAG GCA GCC TTG CTT GTG TGT CAG TAC ACC ATC CAG AGC CTG ATC CAT CTC
met glu ala ala leu leu val cys gln tyr thr ile gln ser leu ile his leu

ACG GGG GTA CCC GGG ACC ATG GGA AAG CCG ATC CCA AAG CCT TTG CTG GGA
thr gly val pro gly thr met gly lys pro ile pro asn pro leu leu gly

TTG GAC TCC ACC TGG CGC GCC GAG AGT
leu asp ser thr trp arg ala glu ser

B)

GAATTC ATG GAG GCA GCC TTG CTT GTG TGT CAG TAC ACC ATC CAG AGC CTG ATC CAT CTC
met glu ala ala leu leu val cys gln tyr thr ile gln ser leu ile his leu

ACG GGG GTA CCT GGG CGC GCC GAG AGT TTA ATA ACC CAA AAG CAA ATC
thr gly val pro gly arg ala glu ser leu ile thr gln lys gln ile

ATG AGC CAA GCA
met ser gln ala

Figure 61: Schematic depicting the DNA & amino acid sequence for the 5' end of the genes encoding the Pk-tagged truncated forms of SV5 HN.

Panel A corresponds to the Δ HNntPk sequence. Panel B represents the Δ HNctPk sequence. Underlined sequences correspond to the putative DNA sequence encoding the EBV gp220/360 cleavable signal sequence. Amino acid sequences in outline text (panel A only) correspond to the 14 amino acid Pk-tag oligopeptide. Amino acids in shadow text correspond to the truncated HN sequence. The ATG codon shown in bold text (panel A) & the underlined ATG shown in shadow text (panel B), represent possible alternative sites for translation initiation.

HN construct. It was not known when the cleavage of the HN-F signal sequence occurred in relation to the completion of synthesis of the polypeptide chain. Therefore, it is not known if dimerisation, and therefore competent transport, is dependent upon attachment to the ER membrane. If it is the case that dimerisation occurs when the two monomers are still attached to the ER membrane, then a lack of secretion could partly be explained by an efficient cleavage of the EBV signal sequence. This may result in a three dimensional diffusion of the monomers to find assembly partners compared with a two dimensional membrane-bound diffusion. This would lead to a slower dimerisation rate of the Δ HN Pk-tagged molecules and therefore inefficient secretion, resulting in the prominent detection of bands of a lower mwt. than expected, due to the lesser extent of post-translational modification. An analysis to determine whether Pk-tagged Δ HN molecules form oligomers has not yet been carried out, but an immune precipitation analysis using the anti-HN mAb 4b would be informative as to whether dimers are formed.

The reason for the presence of the prominent doublet band is unclear, but could be due to several reasons including; degradation of the protein, poor transport of the molecules through the exocytotic pathway (as described above) or, translation initiation from downstream ATGs. The latter explanation is of particular interest since, as described previously, the doublet band was not modified by the addition of N-linked oligosaccharides and therefore has probably not been targeted to the ER. This could be due to the initial ATG (GAATTCATGGAG) occurring in a poor context for translation initiation (particularly sequences 5' to the ATG) as determined by Kozak (1986 & 1989, and detailed in chapter 3 of the introduction). Furthermore, as figure 62 panel B shows, 3' to the putative gp220/360 signal sequence, within the Δ HNctPk sequence, was an ATG codon which is perhaps in a context more suited for translation initiation. If translation initiation occurred at this codon, then the polypeptide would not be directed to the ER, would therefore not be N-linked glycosylated and would thus possess a lower relative mwt. than expected. Furthermore, as with the Δ HNctPk DNA sequence, the Δ HNntPk nucleotide sequence possessed a downstream ATG codon which was in a better context for translation initiation (figure 62 panel A).

An attempt to investigate whether initiation at a downstream ATG was the reason for

the presence of the low mwt. doublet band could not be carried out. This was due to the inability to successfully introduce a more optimal translation initiation context (GCCGCCACCATG) by the insertion of a double stranded oligonucleotide 5' to the EBV-derived signal sequence (data not shown). However, using the constructs available, further studies involving trypsin digest & alkali fractionation of isolated microsomal membranes and *in vitro* transcription/translation (as detailed in Parks & Lamb, 1990), could be performed. These experiments should determine whether the protein is being targeted to the ER, if signal sequence cleavage is occurring, and if the polypeptide is being secreted into the ER lumen.

A truncated form of FctPk was constructed, in which the C-terminal cytoplasmic tail and membrane-anchor were deleted by an appropriate PCR amplification step, resulting in the vector pT7 Δ FctPk. As with the truncated HN constructs, Western blot analysis (figure 35) showed the existence of a prominent doublet band at ~ 60-63 Kda which, in contrast to the Δ HN protein, approximates to the expected molecular weight for glycosylated truncated Pk-tagged F. Longer exposure of the film showed the detection of a further 2 bands at ~58 & 53Kda. Due to their relative molecular weights, the 4 bands appeared to have undergone post-translational modification (unmodified monomeric Pk-tagged truncated F should have a mwt. of ~ 48 Kda when electrophoresed under non-reducing conditions). Indeed, upon digestion with PNGase F, the 4 bands were reduced to a single mwt. species of ~ 48 Kda, thereby indicating that Δ FctPk had undergone N-linked glycosylation (figure 22). The presence of the 4 bands may therefore reflect differential glycosylation of individual polypeptides or, due to a slow transport rate, polypeptides from different compartments of the exocytotic pathway are being detected.

A comparison of the production of Δ FctPk to a truncated form of SV5 F which lacks the C-terminal cytoplasmic tail and transmembrane anchor (Δ F) was not made since a Δ F-expressing plasmid was not constructed. However, a similarly truncated form of the influenza virus HA (also a type I glycoprotein) gene had been constructed by Paterson & Lamb (1987) and shown to produce only a single mwt. species when expressed in CV-1 cells by infection with a recombinant SV40 virus. Whether this reflects a difference in the

properties of the proteins, differences in the detection system, or in the transient expression system utilised, remains to be determined.

Furthermore, as only the ~48 Kda band could be detected when WCEs from SV5 Δ FctPk expressing cells were treated with PNGase F and electrophoresed under reducing conditions, then this would imply that the expressed Δ FctPk molecule has not undergone post-translational cleavage from Δ F₀ to Δ F₁ & F₂. This phenomenon has also been observed for a C-terminal deletion mutant (but not including the membrane-anchor) of MeV F protein where proteolytic cleavage did not occur, and the mutant protein was inefficiently transported to the cell-surface (Sergel & Morrison, 1995). Furthermore, a chimaeric RSV F mutant produced by Anderson *et al.* (1992) which did not form oligomers and was not proteolytically activated was glycosylated - but not to complex, endo H resistant forms. Therefore, it would be interesting to determine whether SV5-derived Δ FctPk molecules are endoH resistant or sensitive, since this would give an indication of the position of the molecules in the exocytotic pathway.

Using chemical cross-linking analysis, Russell *et al.*, (1994) demonstrated that the SV5 F protein probably forms non-covalently linked trimers. Using a similar method, it was demonstrated that, in a small proportion of the expressed molecules, Δ FctPk forms dimers and possibly trimers. This result is in some agreement with Reitter *et al.* (1995) whereby the cytoplasmic tail and the transmembrane anchor of MeV F were deleted, but still resulted in the ability of the truncated F monomers to form trimers. However, in contrast to the results obtained for oligomerisation of SV5-derived Δ FctPk, the majority of the detected truncated MeV F molecules appeared to form oligomers. Whether this difference is due to the sucrose density centrifugation technique used by Reitter *et al.* as a means of oligomer analysis, compared to the chemical cross-linking analysis as used above, remains to be determined.

However, although the Δ FctPk monomer was glycosylated and did form an oligomeric structure in a minority of the expressed molecules, secretion of Pk-tagged Δ F into the medium could not be detected in a Western blot analysis, using the anti-Pk mAb, from samples of medium and concentrated medium (figure 35). This is in contrast to results obtained by Paterson & Lamb, (1987) where a truncated version of influenza virus HA was

secreted into the medium, albeit in small amounts. Furthermore, Reitter *et al.* (1995) demonstrated successful secretion of a similarly truncated version of the MeV F protein into the medium. This lack of detectable secretion of Δ FctPk could be due to cleavage of the Pk-tag occurring, protein adsorbing to the Amicon membrane, protein adsorbing to the cell-surface, or a lack of secretion into the medium. An analysis to determine whether the Pk-tag has been cleaved from the protein could not be determined for reasons mentioned earlier regarding the fragility of the 293 cells/vaccinia virus expression system. An immunoprecipitation analysis performed on the medium to discount the second possibility should be carried out. As with the Δ HN constructs, tryptic digest & alkali fractionation of isolated microsomal membranes could be performed to determine whether Δ FctPk is found as a soluble protein in the lumen of the ER & golgi complex. Furthermore, an attempt was made at performing an immunofluorescence analysis on transfected cells to determine the cellular localisation of Δ FctPk but, due to technical difficulties, the result has not been resolved, and further experimentation is required.

2. / Isolation of cell-lines which produce Pk-tagged HN.

Given the encouraging results from section 1, which demonstrated that the addition of the Pk-tag epitope to the N or C-terminus of HN & C-terminus of F did not significantly alter the processing of the glycoproteins through the mammalian-cell exocytotic pathway, it would be beneficial to isolate cell-lines which expressed the Pk-tagged glycoproteins for protein production & purification, and for experiments to further characterise Pk-tagged glycoprotein processing.

2.1 / Constitutive expression of Pk-tagged HN.

Initial studies to isolate cell-lines which constitutively expressed Pk-tagged HN employed EBV-based vectors. These vectors were chosen because they remain episomal in primate cell nuclei and therefore result in a greater frequency of drug-resistance cell isolation, combined with a moderately high gene copy and therefore gene expression (Yates *et al.*, 1985). Numerous EBV-based vectors were constructed which encoded the HNctPk protein,

where transcription should be initiated at sequences derived from the hCMV major IE enhancer/promoter or the RSV LTR. However, no expression of HNctPk could be detected from transient expression analyses in 293 & 293/E (which constitutively express the EBV EBNA-1 protein) cells, or from drug-resistant cell-lines which were isolated after transfection and selection by drug-resistance. The reason for the lack of expression was unknown and therefore it was decided to construct a set of non-EBV-based vectors (the pcDNA series of vectors), including the gene encoding the HNntPk protein, in which the transcription of the inserted gene could be initiated from either the hCMV promoter, or the T7 ϕ 10 promoter. As expression was observed previously when transcription was driven from the T7 ϕ 10 promoter, then this would act as a positive control. Indeed, expression could be observed, by Western blot analysis, from the pcDNA-based vectors when transcription was initiated from the T7 promoter (figure 39). However, detectable expression was still not observed when transcription was initiated from the hCMV promoter (present in the same plasmid) upon transfection into 293 cells, or into Cos-7 cells where amplification of the transfected plasmid usually results in high levels of protein production (figure 40).

Possible reasons for a lack of detectable expression when transcription was initiated from the hCMV promoter include; poor or non-existent transcription initiation of the gene, inadequate or lack of post-transcriptional processing of the synthesised RNA or, mutations occurring within the gene preventing detection. The latter possibility is unlikely since HNctPk & HNntPk expression was detected when expression was driven from the T7 ϕ 10 promoter, but this explanation cannot be ruled out for the Pk-tagged HN genes present in the EBV-based vectors. Poor or non-existent transcription of the HNctPk gene may be due to a faulty hCMV promoter. However, replacement of the HNctPk gene, within the pcDNA-based plasmid, with the bacterial β -galactosidase gene, resulted in the expression of β -galactosidase from the hCMV promoter, thereby arguing against a faulty promoter being the source for the lack of detectable expression.

Therefore, perhaps the most likely reason for a lack of detectable expression was a failing at the level of RNA processing, perhaps due to the lack of an intron. Indeed, expression of SV5 HN in mammalian cells has, to date, only ever been initiated from an SV40

promoter, as part of an SV40 recombinant virus or by transfection into Cos cells, or via the vaccinia virus/T7 system. Therefore, it is not known whether expression of SV5 HN from the hCMV promoter requires the presence of an appropriate intron for successful expression. A Northern blot analysis should be carried out to determine if transcription is occurring.

Furthermore, others have also extensively tried to isolate cell-lines which constitutively expressed the native SV5 HN protein but have been unsuccessful (R.A. Lamb, personal communication). This apparent inability to construct cell-lines which constitutively expressed SV5 HN may be unique to SV5, since cell-lines which constitutively expressed PIV-4A HN, MuV HN and MeV H have been isolated (Nishio *et al.*, 1994; Hishiyama *et al.*, 1996 and Hirano & Wong, 1991). Therefore, in the assumption that expression of SV5 HN, and therefore Pk-tagged SV5 HN, may be uniquely toxic to the cells, an inducible expression system based upon the tetracycline transactivator protein was employed.

2.2 / Inducible expression of Pk-tagged HN.

Non-EBV-based and EBV-based vectors were constructed in which the HNctPk & HNntPk genes were under the transcriptional control of the tetracycline transactivator protein. Western blot analysis (figures 43 & 51) and immunofluorescence analysis (figure 48) revealed that, in a transient expression analysis, HNntPk was successfully expressed from the various constructs. However, expression of HNctPk could not be detected in a transient expression analysis from any of the constructs. The successful expression of HNntPk from the tetracycline responsive promoter (TRP), would tend to argue against the explicit requirement of an intron for the successful expression of Pk-tagged HN proteins. Furthermore, it is difficult to perceive there exists an absolute requirement for an intron before HNctPk expression occurs whilst no such requirement exists for HNntPk protein production. Therefore, perhaps at least for the HNctPk vectors under control of the TRP, the most likely reason for a lack of expression is due to the presence of mutations which occurred during construction of the plasmid, thereby preventing successful expression of the HNctPk protein. Re-cloning of HNctPk may therefore, prove beneficial.

Despite the disappointing lack of expression of the HNntPk protein, construction of cell-lines which inducible express HNntPk from EBV-based vectors were attempted, due to the successful expression of HNntPk from the tetracycline-responsive promoter in a transient assay. An initial attempt at selecting for drug-resistant 293 cells which had been co-transfected with the constructs pEBV-tTa/neo & pTEP4HNntPk (->->) resulted in only one drug-resistant colony growing up from an initial 10^6 transfected cells. Subsequent analysis of this clone proved negative for both HNntPk & tTa expression.

Since only one drug-resistant clone was produced, two approaches to try and increase this number were examined. The first approach utilised an EBNA-1 expressing 293-derived cell-line, in which the constitutive expression of EBNA-1 facilitates the retention & replication of EBV-based vectors upon transfection, resulting in a greater frequency of drug-resistant colonies. However, the EBNA-1 expressing 293-derivative also expresses neomycin aminoglycoside phosphotransferase. Therefore, the tTa/neo expressing constructs could not be utilised and similar vectors expressing tTa adjoined to the gene coding for hygromycin B phosphotransferase (hyg B PT) were constructed. As figure 46 demonstrates, functional tTa and hyg B PT were expressed from the tTa/2A/hyg B vectors, thereby indicating that the addition of an extra amino acid (proline, derived from the 2A sequence after cleavage) to the N-terminus of hygB PT did not inhibit its function. However, upon transfection and selection of 293/E cells with pEBV-tTa/hygB & pTEP8HNntPk (->->) (in the presence of tetracycline), again only 1 colony grew from an initial 10^6 cells transfected which tested negative for tTa & HNntPk expression by immunofluorescence (data not shown).

However, the second approach in which native 293 cells were transfected with the plasmid pEBV-tTa/neo and subsequently selected with Geneticin, resulted in the successful isolation of a cell-line (table 5) which expressed high-levels of the tTa protein. This cell-line was transfected with the EBV-based HNntPk expressing plasmids pTEP4HNntPk (->->) & pTEP8HNntPk (->->), and are currently undergoing selection for the isolation of drug-resistant cell-lines.

2.3 / Influence of vector type on transient gene expression.

In an attempt to construct cell-lines which inducibly expressed Pk-tagged HN, various driver and responder vectors were constructed in which episomal or non-episomal sequences are present. In a transient expression analysis in EBNA-1 expressing 293 cells, it was observed that the presence of EBV-sequences on the responder and driver plasmids significantly increased the production of HNntPk (figures 51 & 52). The increase in protein production from the EBV-based plasmids is most likely due to their continued retention and replication in the cell nucleus. Furthermore, it must be noted that the observed increase in expression of HNntPk is probably conservative, as equivalent weights of EBV and non-EBV plasmids were transfected. Therefore, the actual number of EBV-based plasmid molecules transfected into the 293/E cells would have been approximately half as many compared to the non-EBV plasmids, due to the larger size of the EBV-based vectors.

3 / Amplification of EBV-based vectors.

High-level recombinant protein production from mammalian cells has routinely been achieved through the amplification of genome-integrated sequences, whereby copy numbers reaching 100 to 1000 are frequently obtained after amplification. However, this process of amplification usually takes place over a long period involving up to 3 months work due to the time required for drug-selection and clonal isolation. A perhaps more attractive approach to gene amplification is through the amplification of episomally maintained plasmids.

The amplification of episomes within mammalian cells is not a novel idea and has been successfully demonstrated by Dubridge *et al.* (1985) in which an SV40 *ori*-containing BPV-based vector, present at a stable copy number in a mouse cell line, was amplified (50-100 fold, to a maximum copy number of ~10,000 per cell) upon fusion with the SV40 LTag-expressing cell-line Cos-7. Furthermore, Haase *et al.* (1989) and Heinzl *et al.* (1988) have demonstrated the amplification of EBV-based vectors in mammalian cells, whereby SV40 *ori*-containing EBV-based plasmids were amplified by the transient production of the SV40 LTag, to a maximum copy number of ~1000 per cell.

Furthermore, an important feature of the SV40 LTA_g amplification system is that there is not a significant increase in the frequency of point mutations which occur in the vector after amplification, although there was a small, but demonstrable, increase in the frequency of deletion mutations, especially if the plasmids contained inverted repeat regions (Heinzel *et al.*, 1988).

Although, as detailed above, amplification of EBV-based vectors has already been established, it has not been investigated whether the increase in the EBV-based plasmid copy number results in a subsequent increase in gene expression & protein production of an encoded protein. Furthermore, the above amplification findings have been performed in cells which stably maintain an episome and are amplified via transfection with plasmids encoding the SV40 LTA_g, or are fused to cells expressing the LTA_g. Neither of these two situations are satisfactory if even a small scale-up in protein production was required.

Therefore, a system was envisaged whereby a primate cell-line would be produced which could inducibly express the SV40 LTA_g, using the tTa system. The tTa scheme was chosen as the inducible system primarily for three reasons; (1) the tTa system has been utilised in this laboratory and the components were immediately available, (2) basal levels of expression in the uninduced phenotype have been demonstrated to be negligible or very low, and (3) ease of induction, by the removal of tetracycline. Clones would be isolated which produced LTA_g to high-levels upon induction, but which demonstrated negligible production in the non-induced phenotype. The cell-line (in an uninduced state) would then be transfected with an SV40 *ori*-containing EBV-based plasmid which encodes the gene of interest under a constitutive promoter such as the hCMV IE enhancer/promoter. Cell-lines which produced the protein of interest to varying levels, would be isolated and the episome amplified by the induction of the SV40 LTA_g and protein production, prior to and after induction of LTA_g, would be determined. Therefore, the first stage of the project was to construct a suitable plasmid encoding the tTa protein. For this purpose the plasmid pGEM-tTa/hygB (which does not possess an SV40 *ori*) was successfully constructed and shown to express functional tTa and hygromycin B phosphotransferase (figures 45 & 46). The second

part of the project was to construct a suitable amplifiable EBV-based plasmid which encoded a non-toxic protein.

3.1 / Production of the SV5 P protein from an SV40 ori-containing EBV-based plasmid.

Two EBV-based vectors which possessed an SV40 *ori* and which encode the non-toxic SV5 P protein downstream from the hCMV enhancer/promoter, were constructed and shown to express SV5 P by immunofluorescence and Western blot analysis (figures 54 and 55). The two plasmids differed in that the orientation of the P transcription unit was opposite to that of the EBNA-1 transcript in the vector designated pCEP8/P (-><-), whilst the vector termed pCEP8/P (->->), had the P transcript in the same orientation to that of EBNA-1. The (->->) vector was shown to produce more P protein in a transient expression assay when compared to the (-><-) vector (figure 56).

This result is in agreement with the findings of Young *et al.* (1988), where the similar orientation of transcripts resulted in more protein production than the opposite orientation. However, in contrast to the results obtained by Young *et al.* where the interferon-gamma gene positioned in the opposite orientation to that of EBNA-1 resulted in negligible production of the protein (compared to the high-level production observed when the transcripts were in the same orientation), there was only a moderate ~3 fold difference in P protein production from the two plasmids (figure 56). This observed difference may be a reflection on the transcription activation elements used by Young *et al.* which consisted of either a; hCMV IE enhancer combined with an SV40 promoter, or a moloney sarcoma virus enhancer combined with a metallothionein promoter, whereas, in the experiment reported here, the P protein was driven by a hCMV IE enhancer/promoter element. A further factor in the observed difference may be due to the use, by Young *et al.*, of monkey CV-1 cells compared to human 293 cells that were utilised in P protein production. However, this moderate difference in protein production between the two orientations was also observed for the HNntPk protein when expression was driven from the tetracycline responsive promoter (figures 51 & 52), which may suggest that the substantial difference in protein production observed by Young *et al.* is the exception rather than the norm.

Furthermore, synchronisation of 293/E cells at the G2/M phase of the cell-cycle prior to transfection, did not result in an increased production of the P protein when compared to unsynchronised cells (figure 57). This is in contrast to the findings of Teshigawara & Katsura (1992), who demonstrated a 10-fold increase in IL-2 receptor β -chain production after transfection of their EBV-based plasmid into G2/M synchronised cells. This increase in protein production was attributable to an increase in plasmid copy number as determined by a Southern blot analysis of a Hirt extract. However, the analysis of protein production carried out by Teshigawara & Katsura was performed on transiently transfected EBV-derived B-cells (Raji) in which transfection efficiencies (determined by fluorescent activated cell sorting) was 24 % in the unsynchronised cell population (90% transfection efficiency in the synchronised cell population), whereas the P protein experiment was analysed on transiently transfected embryonic kidney cells (293), with a transfection efficiency (determined by β -gal. assay) of only ~10%. However, another possible reason for the observed lack of increased protein production after synchronisation, is that perhaps the 293 cells were not in fact synchronised since, other than the visual observation that the cells has slowed in their replication rate, a direct analysis was not performed to confirm synchrony of the population, .

Having successfully constructed appropriate EBV-based plasmids and demonstrated the successful production of a non-toxic protein (the SV5 P protein) encoded from these plasmids, the next part of the project was to construct the plasmid which would inducibly produce the SV40 LTA_g.

3.2 / Production of the SV40 LTA_g protein.

Prior to constructing the plasmid which should inducibly express the LTA_g, the PCR product obtained from the PCR-amplification of the cDNA of the LTA_g gene (from the plasmid pSVori) was sub-cloned into the vector pcDNA1/amp (due to necessary restriction endonuclease sites being available) to produce the plasmid pCT7LTA_g. Transient transfection of the plasmid into 293 cells resulted in the detection of a nuclear-localised protein (figure 59), as expected for the SV40 LTA_g. However, further analysis by Western

blot (figure 60, panel A) revealed the presence of a prominent band at ~ 33 Kda, with minor bands at ~ 40 Kda, 55 Kda and 90 Kda (the latter of which corresponds to full length modified SV40 LTA_g).

It is not clear why the lower mwt. species were detected since similar sized products were not detected from the control Cos-7 cell extracts or from the pSV_{ori} transfected cell extracts. PCR-induced mutagenesis resulting in premature translation-termination might have explained the presence of the lower bands, if it was not for the detection of some full length product. Furthermore, 5 separate transformants demonstrated a similar detection pattern (figure 60, panel B). Perhaps expression of the LTA_g gene from the T7 ϕ 10 promoter in vTF7-3 infected cells might further elucidate why the lower mwt. bands are being detected, i.e. is the lack of an intron affecting the expression of the LTA_g gene. Due to the unexpected truncated product, sub-cloning into the plasmid for inducible expression of LTA_g was not undertaken. Therefore, a cell-line which inducibly expressed the LTA_g was not produced. However, the feasibility of increasing protein production from constitutively expressing cell-lines, using the SV40 LTA_g to amplify SV40 *ori*-containing EBV-based plasmids, was investigated.

3.3 / Analysis of the transient amplification experiment.

For this purpose, 293-based cell-lines were constructed which constitutively produced the SV5 P protein encoded on an EBV-based plasmid. Five cell-lines were characterised and shown to have varying levels of P protein production. However, transient transfection of the SV40 LTA_g plasmid, pSV_{ori}, into the cell-lines did not result in any significant increase in expression of the P protein (figure 61). The lack of an increased level in protein production could be due to the following reasons;

- 1) Insufficient amounts of LTA_g were produced, leading to insufficient plasmid amplification. Indeed experiments performed by Gerard & Gluzman (1985) demonstrated a direct correlation between high-level expression of SV40 LTA_g and high-level replication of SV40 DNA. Furthermore, the expression of the LTA_g gene from the pSV_{ori} plasmid is driven by the SV40 early promoter which has been demonstrated to be down regulated by the

adenovirus E1a protein (Veliech & Ziff, 1985) which is constitutively expressed in 293 cells (Graham *et al.*, 1977). Indeed, as figure 60 panel A demonstrated, the amount of LTA_g produced in the pSV_{ori} transfected 293 cells is considerably less than in Cos-7 cells which expressed high levels of LTA_g.

2) It is possible, although unlikely, that the pCEP8/P (->->) plasmid might in fact not be maintained as an episome and has instead been integrated into the genome. Indeed, Stary & Sarasin (1992) have demonstrated the isolation of a cell clone in which an EBV-based vector had integrated into the Hela cell genome. A Hirt extract (Hirt, 1969) followed by a Southern blot analysis (or PCR analysis) would determine whether this is the case or not.

3) The transfection efficiency of the pSV_{ori} plasmid into the P cell-lines (as determined by a β -gal. assay) was only ~10 %. Therefore, perhaps there were insufficient cells expressing the LTA_g, within the transfected population, to produce a significant increase in protein production.

4) It is also possible that sequences present within the EBV-based vectors are inhibitory to amplification by the SV40 LTA_g. However, Heinzel *et al.* (1988) have demonstrated that EBNA-1 & *ori* P sequences, at least, are not inhibitory for replication by SV40 LTA_g. Transfection of the pCEP8/P vectors into Cos-7 cells would resolve whether inhibitory sequences are present.

5) It is also possible that substantial plasmid amplification did occur, but that there was not a corresponding increase in protein production. An analysis by Hirt extraction followed by Southern blotting would clarify the matter.

4 / Future Developments.

4.1 / Pk-tag & anti-Pk mAb for use as an epitope-tagging system.

It has been demonstrated that the Pk-tag oligopeptide does not interfere with the processing and antigenicity of the SV5 HN & F glycoproteins when attached to the termini of these proteins, and will therefore further the use of Pk-tag for the epitope-tagging of glycoproteins produced in mammalian cells. The above analysis was primarily directed

towards the production of Pk-tagged glycoproteins for use in vaccine development, and therefore the functional characteristics of the Pk-tagged HN & F have not been determined. Therefore, an analysis to determine whether or not the presence Pk-tag at the termini of HN or F interferes with the haemagglutinin & neuraminidase function of HN, and the fusion function of F, should be carried out. Furthermore, the above analysis has only been performed on two model glycoproteins and further analysis with other Pk-tagged glycoproteins expressed in mammalian cells will support the use of Pk-tag/anti-Pk mAb for use as a "universal" epitope-tagging system. Moreover, if problems were encountered when using the 14 amino acid Pk-tag, then the length of the tag can be reduced to 9 amino acids without loss of affinity to the anti-Pk mAb (Southern *et al.*, 1991). Furthermore, preliminary data suggests that the epitope recognised by the anti-Pk mAb may even be as small as 5 amino acids (R.E. Randall, personal communication).

However, a drawback in the Pk-tag/anti-Pk mAb system, in comparison to for example the BTag system, is that the current anti-Pk monoclonal antibody has such a high affinity for the Pk-tag epitope that, for the purpose of immuno-affinity purification, denaturing conditions have to be employed for successful antibody/antigen separation. However, this problem should be rectified by the isolation of lower binding-affinity Pk-tag specific monoclonal antibodies which are currently being produced and screened (C. Dunn, Ph.D project, University of St. Andrews).

4.2 / Novel vectors for recombinant protein production in mammalian cells.

It has been observed by Haase *et al.* (1989) that after amplification of SV40 *ori*-containing plasmids by a transient production of LTA_g, within only 12 days the plasmid copy number had significantly dropped. The aforementioned system of controlled inducible expression of LTA_g would allow "boosting" of the copy number at regular intervals. Furthermore, an adaptation of this system is envisaged where selectable markers such as DHFR or MDR-1 could be added to the EBV-based plasmids, in addition to the current histidinol or hygromycin B resistance markers. Initial expressing cell-lines would be selected using hygB or hisD, and after amplification of the plasmid copy number by transient

production of the LTA_g (either inducibly expressed, or transfected into the cells via plasmid DNA) cells in which the copy number has been boosted to high-levels could be selected for by incubation in high concentrations of selectable drug (such as methotrexate, if DHFR, or colchicine, if MDR-1). Maintenance of the cells in the high concentration of drug may maintain a high episomal copy number. This system would also be useful for the amplification of plasmids encoding toxic proteins, as the protocol would allow the isolation of stable high copy number clones without the requirement for induction of protein production (if the protein was under control of tTa, then induction of LTA_g production would also lead to induced production of the toxic protein).

Many experiments are performed in cell-lines derived from rodents, and subsequently the above systems would be redundant in those cells due to the need for primate-cell factors for the replication of plasmids by EBNA-1 and SV40 LTA_g. However, a suitable alternative episomal amplification system for rodent cells could be conceptualised by employing the rodent-cell-replication-competent-EBV-based-plasmid (Krysan & Calos, 1993) with the addition of a polyomavirus *ori*, whereby expression of polyomavirus LTA_g would lead to the amplification of the episomal plasmid in rodent cells.

An even more interesting and exotic development could be to produce a dual nuclear and cytoplasmic amplification system (NACA system, figure 63), whereby the nuclear amplification stage of a DNA virus is combined with the cytoplasmic amplification phase of an RNA virus. In one scenario, as previously detailed, a cell-line would be produced where the SV40 LTA_g was under tTa control. An SV40 *ori*-containing EBV-based plasmid would be constructed which contained the gene-of-interest adjoined to the replicase/transcriptase (*rep*) coding sequence from the alphavirus: semliki forest virus, in which expression of the transcript is under tTa control. Upon induction of the tetracycline responsive promoter, amplification of the episome in the nucleus occurs along with the concurrent expression of the *rep*/gene-of-interest transcript. The *rep*/gene-of-interest transcript is transported to the cytoplasm whereby translation of the mRNA results in the production of the *rep* genes and the subsequent high-level amplification of the *rep*/gene-of-interest transcript followed by a simultaneous high-level expression of the gene-of-interest. The firefly luciferase protein or

the green fluorescent protein would be suitable reporter proteins for initial analysis of the feasibility of the system.

Furthermore, it has been observed by Haase *et al.* (1994) that high-level transcription can inhibit the replication of autonomously replicating plasmids. Therefore, if this was found to be a problem, the above system could be modified by, as previously described, adding selectable markers encoding DHFR or MDR-1 to the plasmid vector, and selecting for cells harbouring high copy number episomes after the transient introduction of the LTA_g. Expression of the rep/gene-of-interest should not occur since the cells would be in the uninduced state, and therefore transcription should not interfere with the plasmid amplification phase of the system.

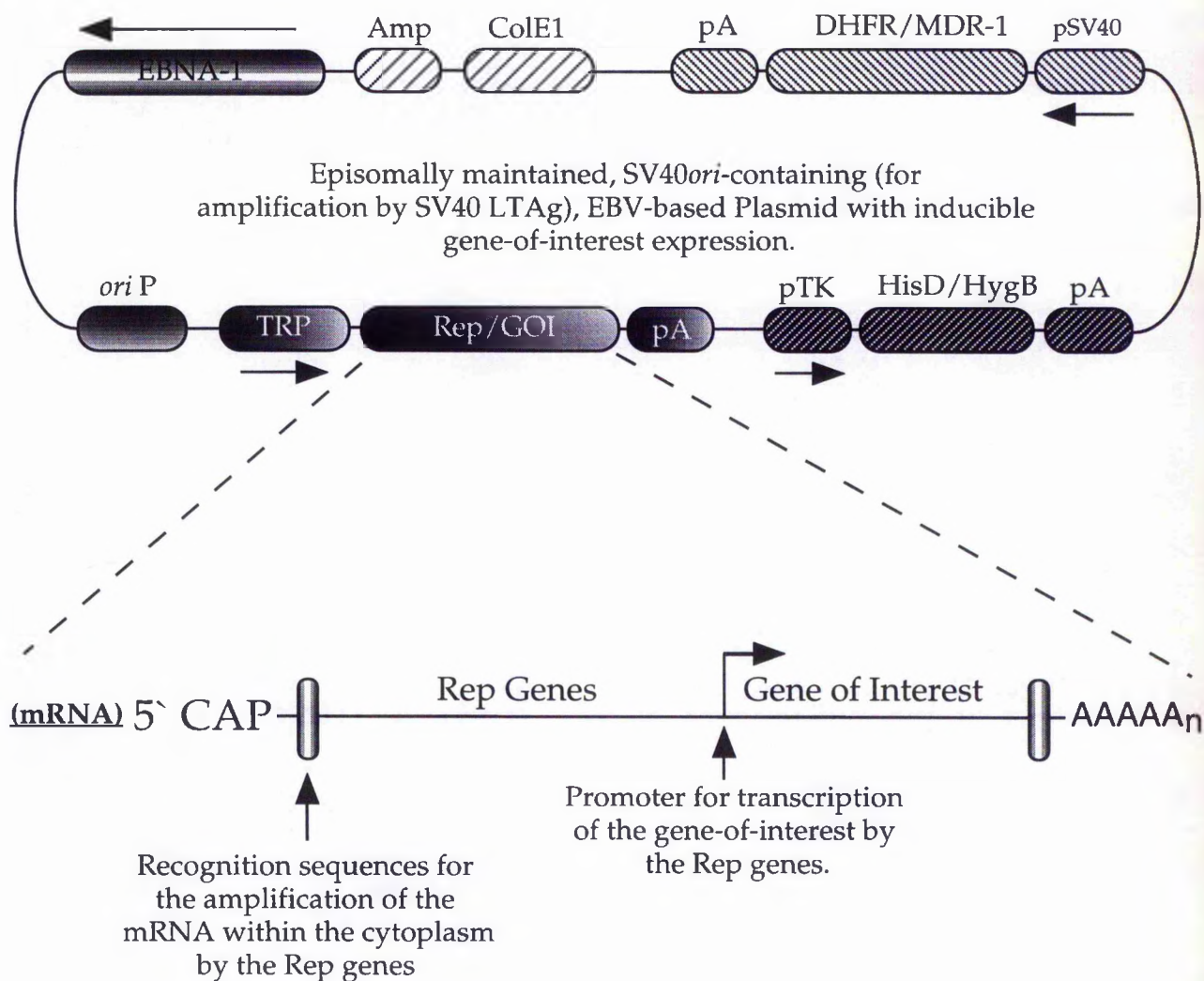


Figure 62: Diagrammatic representation of the components involved in the proposed scheme of a dual Nuclear And Cytoplasmic Amplification (NACA) system.

The depicted plasmid is transfected into primate cells which constitutively express high-levels of the tTa driver protein, and drug resistant cells are isolated using histidinol or hygromycin B. The episomal-containing cell-line is then transfected with a plasmid expressing the SV40 LTA_g, and cells which have amplified the episomal copy number are selected with drugs selecting for high levels of DHFR or MDR-1. Induction of expression of the Rep/GOI transcript leads to high level production of mRNA which is transported to the cytoplasm. Translation of the mRNA leads to synthesis of the Rep proteins which further amplify the mRNA in the cytoplasm and also transcribes & caps the mRNA for the gene-of-interest. Such a process should lead to the production of high-levels of specified protein.

APPENDIX A

Open reading frame of the SV5 HN-gene (W3a).

(1698bp, 565 amino acids, 62,164 Da unmodified).

ATG	GTT	GCA	GAA	GAT	GCC	CCT	GTT	AGG	GCC	ACT	TGC	CGA	GTA	TTA	TTT
Met	val	ala	glu	asp	ala	pro	val	arg	ala	thr	cys	arg	val	leu	phe
CGA	ACA	ACA	ACT	TTA	ATC	TTT	CTA	TGC	ACA	CTA	CTA	GCA	TTA	AGC	ATC
arg	thr	thr	thr	leu	ile	phe	leu	cys	thr	leu	leu	ala	leu	ser	ile
TCT	ATC	CTT	TAT	GAG	AGT	TTA	ATA	ACC	CAA	AAG	CAA	ATC	ATG	AGC	CAA
ser	ile	leu	tyr	glu	ser	leu	ile	thr	gln	lys	gln	ile	met	ser	gln
GCA	GGC	TCA	ACT	GGA	TCT	AAT	TCT	GGA	TTA	GGA	AGT	ATC	ACT	GAT	CTT
ala	gly	ser	thr	gly	ser	asn	ser	gly	leu	gly	ser	ile	thr	asp	leu
CTT	AAT	AAT	ATT	CTC	TCT	GTC	GCA	AAT	CAG	ATT	ATA	TAT	AAC	TCT	GCA
leu	asn	asn	ile	leu	ser	val	ala	asn	gln	ile	ile	tyr	asn	ser	ala
GTC	GCT	CTA	CCT	CTA	CAA	TTG	GAC	ACT	CTT	GAA	TCA	ACA	CTC	CTT	ACA
val	ala	leu	pro	leu	gln	leu	asp	thr	leu	glu	ser	thr	leu	leu	thr
GCC	ATT	AAG	TCT	CTT	CAA	ACC	AGT	GAC	AAG	CTA	GAA	CAG	AAC	TGC	TCG
ala	ile	lys	ser	leu	gln	thr	ser	asp	lys	leu	glu	gln	asn	cys	ser
TGG	AGT	GCT	GCA	CTG	ATT	AAT	GAT	AAT	AGA	TAC	ATT	AAT	GGC	ATC	AAT
trp	ser	ala	ala	leu	ile	asn	asp	asn	arg	tyr	ile	asn	gly	ile	asn
CAG	TTC	TAT	TTT	TCA	ATT	GCT	GAG	GGT	CGC	AAT	CTG	ACA	CTT	GGC	CCA
gln	phe	tyr	phe	ser	ile	ala	glu	gly	arg	asn	leu	thr	leu	gly	pro
CTT	CTT	AAT	ATG	CCT	AGT	TTC	ATT	CCA	ACT	GCC	ACG	ACA	CCA	GAG	GGC
leu	leu	asn	met	pro	ser	phe	ile	pro	thr	ala	thr	thr	pro	glu	gly
TGC	ACC	AGG	ATC	CCA	TCA	TTC	TCG	CTC	ACT	AAG	ACA	CAC	TGG	TGT	TAT
cys	thr	arg	ile	pro	ser	phe	ser	leu	thr	lys	thr	his	trp	cys	tyr
ACA	CAC	AAT	GTT	ATC	CTG	AAT	GGA	TGC	CAG	GAT	CAT	GTA	TCC	TCA	AAT
thr	his	asn	val	ile	leu	asn	gly	cys	gln	asp	his	val	ser	ser	asn
CAA	TTT	GTT	TCC	ATG	GGA	ATC	ATT	GAA	CCC	ACT	TCT	GCC	GGG	TTT	CCA
gln	phe	val	ser	met	gly	ile	ile	glu	pro	thr	ser	ala	gly	phe	pro
TTC	TTT	CGA	ACC	CTA	AAG	ACT	CTA	TAT	CTC	AGC	GAT	GGG	GTC	AAT	CGT
phe	phe	arg	thr	leu	lys	thr	leu	tyr	leu	ser	asp	gly	val	asn	arg
AAG	AGC	TGC	TCT	ATC	AGT	ACA	GTT	CCG	GGG	GGT	TGT	ATG	ATG	TAC	TGT
lys	ser	cys	ser	ile	ser	thr	val	pro	gly	gly	cys	met	met	tyr	cys
TTT	GTT	TCT	ACT	CAA	CCA	GAG	AGG	GAT	GAC	TAC	TTT	TCT	GCC	GCT	CCT
phe	val	ser	thr	gln	pro	glu	arg	asp	asp	tyr	phe	ser	ala	ala	pro
CCA	GAA	CAA	CGA	ATT	ATT	ATA	ATG	TAC	TAT	AAT	GAT	ACA	ATC	GTG	GAG
pro	glu	gln	arg	ile	ile	ile	met	tyr	tyr	asn	asp	thr	ile	val	glu
CGC	ATA	ATT	AAT	CCA	CCC	GGG	GTA	CTA	GAT	GTA	TGG	GCA	ACA	TTG	AAC
arg	ile	ile	asn	pro	pro	gly	val	leu	asp	val	trp	ala	thr	leu	asn
CCA	GGA	ACA	GGA	AGC	GGG	GTA	TAT	TAT	TTA	GGT	TGG	GTG	CTC	TTT	CCA

pro gly thr gly ser gly val tyr tyr leu gly trp val leu phe pro
 ATA TAT GGC GGC GTG ATT AAA GGT ACG AGT TTA TGG AAT AAT CAA GCA
 ile tyr gly gly val ile lys gly thr ser leu trp asn asn gln ala
 AAT AAA TAC TTT ATC CCC CAG ATG GTT GCT GCT CTC TGC TCA CAA AAC
 asn lys tyr phe ile pro gln met val ala ala leu cys ser gln asn
 CAG GCA ACT CAA GTC CAA AAT GCT AAG TCA TCA TAC TAT AGC AGC TGG
 gln ala thr gln val gln asn ala lys ser ser tyr tyr ser ser trp
 TTT GGC AAT CGA ATG ATT CAG TCT GGG ATC CTG GCA TGT CCT CTT CGA
 phe gly asn arg met ile gln ser gly ile leu ala cys pro leu arg
 CAG GAT CTA ACC AAT GAG TGT TTA GTT CTG CCC TTT TCT AAT GAT CAG
 gln asp leu thr asn glu cys leu val leu pro phe ser asn asp gln
 GTG CTT ATG GGT GCT GAA GGG AGA TTA TAC ATG TAT GGT GAC TCG GTG
 val leu met gly ala glu gly arg leu tyr met tyr gly asp ser val
 TAT TAC TAT CAA AGA AGC AAT AGT TGG TGG CCT ATG ACC ATG CTG TAT
 tyr tyr tyr gln arg ser asn ser trp trp pro met thr met leu tyr
 AAG GTA ACC ATA ACA TTC ACT AAT GGT CAG CCA TCT GCT ATA TCA GCT
 lys val thr ile thr phe thr asn gly gln pro ser ala ile ser ala
 CAG AAT GTG CCC ACA CAG CAG GTC CCT AGA CCT GGG ACA GGA GAC TGC
 gln asn val pro thr gln gln val pro arg pro gly thr gly asp cys
 TCT GCA ACC AAT AGA TGT CCC GGT TTT TGC TTG ACA GGA GTG TAT GCC
 ser ala thr asn arg cys pro gly phe cys leu thr gly val tyr ala
 GAT GCC TGG TTA CTG ACC AAC CCT TCG TCT ACC AGT ACA TTT GGA TCA
 asp ala trp leu leu thr asn pro ser ser thr ser thr phe gly ser
 GAA GCA ACC TTC ACT GGT TCT TAT CTC AAC ACA GCA ACT CAG CGT ATC
 glu ala thr phe thr gly ser tyr leu asn thr ala thr gln arg ile
 AAT CCG ACG ATG TAT ATC GCG AAC AAC ACA CAG ATC ATA AGC TCA CAG
 asn pro thr met tyr ile ala asn asn thr gln ile ile ser ser gln
 CAA TTT GGA TCA AGC GGT CAA GAA GCA GCA TAT GGC CAC ACA ACC TGT
 gln phe gly ser ser gly gln glu ala ala tyr gly his thr thr cys
 TTT AGG GAC ACA GGC TCT GTT ATG GTA TAC TGT ATC TAT ATT ATT GAA
 phe arg asp thr gly ser val met val tyr cys ile tyr ile ile glu
 TTG TCC TCA TCT CTC TTA GGA CAA TTT CAG ATT GTC CCA TTT ATC CGT
 leu ser ser ser leu leu gly gln phe gln ile val pro phe ile arg
 CAG GTG ACA CTA TCC
 gln val thr leu ser

APPENDIX B

Open reading frame of the SV5 F-gene (W3a).

(1587bp, 529 amino acids 56,560 Da unmodified).

ATG	GGT	ACT	ATA	ATT	CAA	TTT	CTG	GTG	GTC	TCC	TGT	CTA	TTG	GCA	GGA
Met	gly	thr	ile	ile	gln	phe	leu	val	val	ser	cys	leu	leu	ala	gly
GCA	GGC	AGC	CTT	GAT	CCA	GCA	GCC	CTC	ATG	CAA	ATC	GGT	GTC	ATT	CCA
ala	gly	ser	leu	asp	pro	ala	ala	leu	met	gln	ile	gly	val	ile	pro
ACA	AAT	GTC	CGG	CAA	CTT	ATG	TAT	TAT	ACT	GAG	GCC	TCA	TCA	GCA	TTC
thr	asn	val	arg	gln	leu	met	tyr	tyr	thr	glu	ala	ser	ser	ala	phe
ATT	GTT	GTG	AAG	TTA	ATG	CCT	ACA	ATT	GAC	TCG	CCG	ATT	AGT	GGA	TGT
ile	val	val	lys	leu	met	pro	thr	ile	asp	ser	pro	ile	ser	gly	cys
AAT	ATA	ACA	TCA	ATT	TCA	AGC	TAT	AAT	GCA	ACA	GTG	ACA	AAA	CTC	CTA
asn	ile	thr	ser	ile	ser	ser	tyr	asn	ala	thr	val	thr	lys	leu	leu
CAG	CCG	ATC	GGT	GAG	AAT	TTG	GAG	ACG	ATT	AGG	AAC	CAG	TTG	ATT	CCA
gln	pro	ile	gly	glu	asn	leu	glu	thr	ile	arg	asn	gln	leu	ile	pro
ACT	CGG	AGG	AGA	CGC	CGG	TTT	GCA	GGG	GTG	GTG	ATT	GGA	TTA	GCT	GCA
thr	arg	arg	arg	arg	arg	phe	ala	gly	val	val	ile	gly	leu	ala	ala
TTA	GGA	GTA	GCT	ACT	GCC	GCA	CAG	GTC	ACT	GCC	GCA	GTG	GCA	CTA	GTA
leu	gly	val	ala	thr	ala	ala	gln	val	thr	ala	ala	val	ala	leu	val
AAG	GCA	AAT	GAA	AAT	GCT	GCG	GCT	ATA	CTC	AAT	CTC	AAA	AAT	GCA	ATC
lys	ala	asn	glu	asn	ala	ala	ala	ile	leu	asn	leu	lys	asn	ala	ile
CAA	AAA	ACA	AAT	GCA	GCA	GTT	GCA	GAT	GTG	GTC	CAG	GCC	ACA	CAA	TCA
gln	lys	thr	asn	ala	ala	val	ala	asp	val	val	gln	ala	thr	gln	ser
CTA	GGA	ACG	GCA	GTT	CAA	GCA	GTT	CAA	GAT	CAC	ATA	AAC	AGT	GTG	GTA
leu	gly	thr	ala	val	gln	ala	val	gln	asp	his	ile	asn	ser	val	val
AGT	CCA	GCA	ATT	ACA	GCA	GCC	AAT	TGT	AAG	GCC	CAA	GAT	GCT	ATC	ATT
ser	pro	ala	ile	thr	ala	ala	asn	cys	lys	ala	gln	asp	ala	ile	ile
GGC	TCA	ATC	CTC	AAT	CTC	TAT	TTG	ACC	GAG	TTG	ACA	ACC	ATC	TTC	CAC
gly	ser	ile	leu	asn	leu	tyr	leu	thr	glu	leu	thr	thr	ile	phe	his
AAT	CAA	ATT	ACA	AAC	CCT	GCA	TTG	AGT	CCC	ATT	ACA	ATT	CAA	GCT	TTA
asn	gln	ile	thr	asn	pro	ala	leu	ser	pro	ile	thr	ile	gln	ala	leu
AGG	ATC	CTA	CTG	GGG	AGT	ACC	TTG	CCG	ACT	GTG	GTC	GAA	AAA	TCT	TTC
arg	ile	leu	leu	gly	ser	thr	leu	pro	thr	val	val	glu	lys	ser	phe
AAT	ACC	CAG	ATA	AGT	GCA	GCT	GAG	CTT	CTC	TCA	TCA	GGG	TTA	TTG	ACA
asn	thr	gln	ile	ser	ala	ala	glu	leu	leu	ser	ser	gly	leu	leu	thr
GGC	CAG	ATT	GTG	GGA	TTA	GAT	TTG	ACC	TAT	ATG	CAG	ATG	GTC	ATA	AAA
gly	gln	ile	val	gly	leu	asp	leu	thr	tyr	met	gln	met	val	ile	lys
ATT	GAG	CTG	CCA	ACT	TTA	ACT	GTA	CAA	CCT	GCA	ACC	CAG	ATC	ATA	GAT
ile	glu	leu	pro	thr	leu	thr	val	gln	pro	ala	thr	gln	ile	ile	asp

CTG	GCC	ACC	ATT	TCT	GCA	TTC	ATT	AAC	AAT	CAA	GAA	GTC	ATG	GCC	CAA
leu	ala	thr	ile	ser	ala	phe	ile	asn	asn	gln	glu	val	met	ala	gln
TTA	CCA	ACA	CGT	GTT	ATG	GTG	ACT	GGC	AGC	TTG	ATC	CAA	GCC	TAT	CCC
leu	pro	thr	arg	val	met	val	thr	gly	ser	leu	ile	gln	ala	tyr	pro
GCA	TCG	CAA	TGC	ACC	ATT	ACA	CCC	AAC	ACT	GTG	TAC	TGT	AGG	TAT	AAT
ala	ser	gln	cys	thr	ile	thr	pro	asn	thr	val	tyr	cys	arg	tyr	asn
GAT	GCC	CAA	GTA	CTC	TCA	GAT	GAT	ACT	ATG	GCT	TGC	CTC	CAA	GGT	AAC
asp	ala	gln	val	leu	ser	asp	asp	thr	met	ala	cys	leu	gln	gly	asn
TTG	ACA	AGA	TGC	ACC	TTC	TCT	CCA	GTG	GTT	GGG	AGC	TTT	CTC	ACT	CGA
leu	thr	arg	cys	thr	phe	ser	pro	val	val	gly	ser	phe	leu	thr	arg
TTC	GTG	CTG	TTC	GAT	GGA	ATA	GTT	TAT	GCA	AAT	TGC	AGG	TCG	ATG	TTG
phe	val	leu	phe	asp	gly	ile	val	tyr	ala	asn	cys	arg	ser	met	leu
TGC	AAG	TGC	ATG	CAA	CCT	GCT	GCT	GTG	ATC	CTA	CAG	CCG	AGT	TCA	TCC
cys	lys	cys	met	gln	pro	ala	ala	val	ile	leu	gln	pro	ser	ser	ser
CCT	GTA	ACT	GTC	ATT	GAC	ATG	TAC	AAA	TGT	GTG	AGT	CTG	CAG	CTT	GAC
pro	val	thr	val	ile	asp	met	tyr	lys	cys	val	ser	leu	gln	leu	asp
AAT	CTC	AGA	TTC	ACC	ATC	ACT	CAA	TTG	GCC	AAT	GTA	ACC	TAC	AAT	AGC
asn	leu	arg	phe	thr	ile	thr	gln	leu	ala	asn	val	thr	tyr	asn	ser
ACC	ATC	AAG	CTT	GAA	TCA	TCC	CAG	ATC	TTG	TCT	ATT	GAT	CCG	TTG	GAT
thr	ile	lys	leu	glu	ser	ser	gln	ile	leu	ser	ile	asp	pro	leu	asp
ATA	TCC	CAA	AAT	CTA	GCT	GCG	GTG	AAT	AAG	AGT	CTA	AGT	GAT	GCA	CTA
ile	ser	gln	asn	leu	ala	ala	val	asn	lys	ser	leu	ser	asp	ala	leu
CAA	CAC	TTA	GCA	CAA	AGT	GAC	ACA	TAT	CTT	TCT	GCA	ATC	ACA	TCA	GCT
gln	his	leu	ala	gln	ser	asp	thr	tyr	leu	ser	ala	ile	thr	ser	ala
ACG	ACT	ACA	AGT	GTA	TTA	TCC	ATA	ATA	GCA	ATC	TGT	CTT	GGA	TCG	TTA
thr	thr	thr	ser	val	leu	ser	ile	ile	ala	ile	cys	leu	gly	ser	leu
GGT	TTA	ATA	TTA	ATA	ATC	TTG	CTC	AGT	GTA	GTT	GTG	TGG	AAG	TTA	TTG
gly	leu	ile	leu	ile	ile	leu	leu	ser	val	val	val	trp	lys	leu	leu
ACC	ATT	GTC	GTT	GCT	AAT	CGA	AAT	AGA	ATG	GAG	AAT	TTT	GTT	TAT	CAT
thr	ile	val	val	ala	asn	arg	asn	arg	met	glu	asn	phe	val	tyr	his
AAA	TAA														
lys	OCH														

REFERENCES

- ABDUL JABBAR, M. and NAYAK, D.P. (1987). *Signal processing, glycosylation and secretion of mutant haemagglutinins of a human influenza virus by S. cerevisiae*. MOL. AND CELL. BIOL. 7(4):1476-1485.
- AKUSJARVI, G., SVENSSON, C. and NYGARO, O. (1987). *A mechanism by which adenovirus virus-associated RNA₁ controls translation in a transient expression assay*. MOL. AND CELL. BIOL. 7(1):549-551.
- ANDERSON, K., STOTT, E.J. and WERTZ, G.W. (1992). *Intracellular processing of the human respiratory syncytial virus fusion glycoprotein: amino acid substitutions affecting folding, transport and cleavage*. J. GEN. VIROL. 73:1177-1188.
- ANTON, M. and GRAHAM, F.L. (1995). *Site-specific recombination mediated by an adenovirus vector expressing the Cre recombinase protein: a molecular switch for control of gene expression*. J. VIROL. 69(8):4600-4606.
- BAGAI, S. and LAMB, R.A. (1995). *Quantitative measurement of paramyxovirus fusion: differences in requirements of glycoproteins between simian virus 5 and human parainfluenza virus 3 or newcastle disease virus*. J. VIROL. 69(11):6712-6719.
- BAGAI, S. and LAMB, R.A. (1995). *Individual roles of N-linked oligosaccharide chains in intracellular transport of the paramyxovirus SV5 fusion protein*. VIROLOGY 209(1):250-256.
- BERGERON, J.J.M., BRENNER, M.B., THOMAS, D.Y. and WILLIAMS, D.B. (1994). *Calnexin: a membrane-bound chaperone of the endoplasmic reticulum*. TIBS 19:124-128.
- BERGLUND, P., TUBULEKAS, I. and LILJESTROM, P. (1996). *Alphaviruses as vectors for gene delivery*. TIBTECH 14:130-134.
- BERNS, K.I. *Parvoviridae: The viruses and their replication*. In: *Fields Virology*, edited by Fields, B.N., Knipe, D.M. and Howley, P.M. Philadelphia: Lippincott-Raven, 1996, p. 2173-2197.
- BERTHOLET, C., DRILLIEN, R. and WITTEK, R. (1985). *One hundred base pairs of 5' flanking sequence of a vaccinia virus late gene are sufficient to temporally regulate late transcription*. PROC. NATL. ACAD. SCI. USA 82:2096-2100.
- BLOBEL, G. and Sabatini, D.D. *Biomembranes*, Plenum Publishing, 1971. pp. 193-195.
- BLUMBERG, B.M., GIORGI, C. and KOLAKOFSKY, D. (1983). *N-protein of vesicular stomatitis virus selectively encapsidates leader RNA in vitro*. CELL 32:559-567.
- BOECK, R. and KOLAKOFSKY, D. (1994). *Positions +5 and +6 can be major determinants of the efficiency of non-AUG codons for protein synthesis*. EMBO J 13(15):3608-3617.
- BOHN, W., RUTTER, G., HOHENBERG, H., MANNWEILER, K., et al. (1986). *Involvement of actin filaments in budding of measles virus: studies on cytoskeletons of infected cells*. VIROLOGY 149:91-106.

- BOMSEL, M. and MOSTOV, K. (1991). *Sorting of plasma membrane proteins in epithelial cells*. CURR. OPIN. CELL BIOL. 3:647-653.
- BONIFACINO, J.S. and LIPPINCOTT-SCHWARTZ, J. (1991). *Degradation of proteins within the endoplasmic reticulum*. CURR. OPIN. CELL BIOL. 3:592-600.
- BOUSSE, T., TAKIMOTO, T. and PORTNER, A. (1995). *A single amino acid change enhances the fusion promotion activity of human parainfluenza virus type 1*. VIROLOGY 209:654-657.
- BRICKELL, P.M. and PATEL, M.S. (1995). *Epstein-Barr virus replication studies and their application to vector design*. MOL. BIOTECH. 3:199-205.
- BRINSTER, R.L., ALLEN, J.M., BEHRINGER, R.R., GELINA, R.E., et al. (1988). *Introns increase transcriptional efficiency in transgenic mice*. PROC. NATL. ACAD. SCI. USA 85(3):836-840.
- BUCHHOLZ, C.J., SPEHNER, D., DRILLIEN, R., NEUBERT, W.J., et al. (1993). *The conserved N-terminal region of Sendai virus nucleocapsid protein, NP, is required for nucleocapsid assembly*. J. VIROL. 67:5803-5812.
- BUCHMAN, A.R. and BERG, P. (1988). *Comparison of intron-dependent and intron-independent gene expression*. MOL. AND CELL. BIOL. 8(10):4395-4405.
- CARAVOKYRI, C. and LEPPARD, K.N. (1995). *Constitutive episomal expression of polypeptide IX (pIX) in a 293-based cell line complements the deficiency of pIX mutant adenovirus type 5*. J. VIROL. 69(11):6627-6633.
- CHAKRABARTI, S., BRECHLING, K. and MOSS, B. (1985). *Vaccinia virus expression vector: Co-expression of b-galactosidase gene provides visual screening of recombinant viral plaques*. MOL. AND CELL. BIOL. 5:3403-3409.
- CHARTIER, C., DEGRYSE, E., GANTZER, M., DIETERLE, A., et al. (1996). *Efficient generation of recombinant adenovirus vectors by homologous recombination in Eschericia coli*. J. VIROL. 70(7):4805-4810.
- CHEN, X., LI, Y., XIONG, K. and WAGNER, T.E. (1994). *A self-initiating eukaryotic transient gene expression system based on co-transfection of bacteriophage T7 RNA polymerase and DNA vectors containing a T7 autogene*. NAR 22(11):2114-2120.
- CHIDGEY, M.A.J. (1993). *Protein targeting to dense core secretory granules*. BIOESSAYS 15(5):317-321.
- CHOPPIN, P.W. AND COMPANS, R.W. *Reproduction of Paramyxoviruses*. In: *Comprehensive Virology Volume 4*, edited by Fraenkel-Conrat, H. and Wagner, R.R. New York: Plenum Publishing Corporation, 1975, p. 95-178.
- COLLEY, K.J., LEE, E.U., ADLER, B., BROWNE, J.K., et al. (1989). *Conversion of a Golgi apparatus sialyltransferase to a secretory protein by replacement of the NH2-terminal signal anchor with a signal peptide*. J. BIOL. CHEM. 264(30):17619-17622.
- COLLINS, P.L. *The molecular biology of human respiratory syncytial virus (RSV) of the genus Pneumovirus*. In: *The Paramyxoviruses*, edited by Kingsbury, .. New York: Plenum Publishing Corporation, 1991, p. 103-162.
- COLLINS, P.L. and MOTTET, G. (1991). *Posttranslational processing and oligomerisation of the fusion glycoprotein of human respiratory syncytial virus*. J. GEN. VIROL. 72:3095-3101.

- COLMAN, P.M., VARGHESE, J.N. and LAVER, W.G. (1983). *Structure of the catalytic and antigenic sites in influenza-virus neuraminidase*. NATURE 303:41-44.
- COLMAN, P.M., HOYNE, P.A. and LAWRENCE, M.C. (1993). *Sequence and structure alignment of paramyxovirus hemagglutinin-neuraminidase with influenza virus neuraminidase*. J. VIROL. 67:2972-2980.
- COSSON, P. and LETOURNEUR, F. (1994). *Coatamer interaction with di-lysine endoplasmic reticulum retention motifs*. SCIENCE 263(5153):1629-1631.
- COTTEN, M., WAGNER, E., ZATLOUKAI, K. and BIRNSTIEL, M.L. (1993). *Chicken adenovirus (CELO virus) particles augment receptor-mediated DNA delivery to mammalian cells and yield exceptional levels of stable transformants*. J. VIROL. 67(7):3777-3785.
- CREGG, J.M. (1993). *Recent advances in the expression of foreign genes in Pichia pastoris*. BIOTECH 11:905-910.
- CROWLEY, K.S., REINHART, G.D. and JOHNSON, A.E. (1993). *The signal sequence moves through a ribosomal tunnel into a noncytoplasmic aqueous environment at the ER membrane early in translocation*. CELL 73:1101-1115.
- CURRAN, J., HOMANN, H., BUCHHOLZ, C., ROCHAT, S., et al. (1993). *The hypervariable c-terminal tail of the sendai paramyxovirus nucleocapsid protein is required for template function but not for RNA encapsidation*. J. VIROL. 67:4358-4364.
- CURRAN, J., MARQ, J. and KOLAKOFSKY, D. (1992). *The Sendai virus nonstructural C proteins specifically inhibit mRNA synthesis*. VIROLOGY 189:647-656.
- CURRAN, J., PELET, T. and KOLAKOFSKY, D. (1994). *An acidic activation like domain of the Sendai virus P protein is required for RNA synthesis and encapsidation*. VIROLOGY 202:875-884.
- CZARNOTA, G.J., ANDREWS, D.W., FARROW, N.A. and OTTENSMEYER, F.P. (1994). *A structure for the signal sequence binding-protein SRP54 - 3D reconstruction from stem images of single molecules*. J. STRUCT. BIOL. 113(1):35-46.
- DAHLER, A., WADE, R.P., MUSCAT, G.E.O. and WATERS, M.J. (1994). *Expression vectors encoding human growth hormone (hgh) controlled by human muscle-specific promoters: prospects for regulated production of hgh by myoblast transfer or intravenous injection*. GENE 145:305-310.
- DANKO, I. and WOLFF, J.A. (1994). *Direct gene transfer into muscle*. VACCINE 12(16):1499-1502.
- DAVIS, H.L., MICHEL, M-L., MANCINI, M., SCHLEEF, M., et al. (1994). *Direct gene transfer in skeletal muscle: plasmid DNA-based immunization against the hepatitis B virus surface antigen*. VACCINE 12(16):1503-1509.
- DAVIS, H.L., SCHRIMBECK, R., REIMANN, J. and WHALEN, R.G. (1995). *DNA-mediated immunization in mice induces a potent MHC class I-restricted cytotoxic T-lymphocyte response to the hepatitis-B envelope protein*. HUMAN GENE THERAPY 6(11):1447-1456.
- DENG, R., WANG, Z., MIRZA, A.M. and IORIO, R.M. (1995). *Localization of a domain on the paramyxovirus attachment protein required for the promotion of cellular fusion by its homologous fusion protein spike*. VIROLOGY 209:457-469.
- DEUSCHLE, U., MEYER, W.K.H. and THIESEN, H-J. (1995). *Tetracycline-reversible silencing of eukaryotic promoters*. MOL. AND CELL. BIOL. 15(4):1907-1914.
- DIDCOCK, L.J. Honours project thesis, University of St. Andrews, 1996.

- DOMS, R.W., KELLER, D.S., HELENIUS, A.H. and BALCH, W.E. (1987). Role for ATP in regulating the assembly and transport of vesicular stomatitis virus G protein trimers. *J. CELL. BIOL.* **105**:1957-1969.
- DU, R-P., GAIL, E.D., WYDE, P.R., YAN, W-Y., et al. (1994). A prototype recombinant vaccine against respiratory syncytial virus and parainfluenza virus type 3. *BIOTECH* **12**:813-818.
- DUBENSKY, T.W., DRIVER, D.A., POLO, J.M., BELLI, B.A., et al. (1996). Sindbis virus DNA-based expression vectors: Utility for in vitro and in vivo gene transfer. *J. VIROL.* **70**(1):508-519.
- DUBRIDGE, R.B., LUSKY, M., BOTCHAN, M.R. and CALOS, M.P. (1985). Amplification of a bovine papillomavirus-simian virus 40 chimera. *J. VIROL.* **56**:625-627.
- ECTOR, K.J., NG, D.T.W. and LAMB, R.A. (1992). Specific residues flanking the transmembrane domain of a viral glycoprotein are involved in internalization and lysosomal transport. *MOL. BIOL. CELL* **3**(5):60.
- EDWARDS, C.P. and ARUFFO, A. (1993). Current applications of Cos cell based transient expression systems. *CURR. OPIN. IN BIOTECH.* **4**:558-563.
- ELROY-STEIN, O., FUERST, T.R. and MOSS, B. (1989). Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *PROC. NATL. ACAD. SCI. USA* **86**:6126-6130.
- FIEDLER, K. and SIMONS, K. (1995). The role of N-glycans in the secretory pathway. *CELL* **81**:309-312.
- FINK, D.J., DELUCA, N.A., GOINS, W.F. and GLORIOSO, J.C. (1996). Gene transfer to neurons using herpes simplex virus-based vectors. *ANNUAL REVIEWS IN NEUROSCIENCE* **19**:265-287.
- FUERST, T.R., EARL, P.L. and MOSS, B. (1987). Use of a hybrid vaccinia virus-T7 RNA polymerase system for expression of target genes. *MOL. AND CELL. BIOL.* **7**:2538-2544.
- FUERST, T.R., NILES, E.G., STUDIER, F.W. and MOSS, B. (1986). Eukaryotic transient expression system based upon recombinant vaccinia virus that synthesises bacteriophage T7 RNA polymerase. *PROC. NATL. ACAD. SCI. USA* **83**:8122-8126.
- FULLER, D.H. and HAYNES, J.R. (1994). A qualitative progression in HIV type-1 glycoprotein-120-specific cytotoxic cellular and humoral immune responses in mice receiving a DNA-based glycoprotein-120 vaccine. *AIDS RESEARCH AND HUMAN RETROVIRUSES* **10**(11):1433-1441.
- FUKUSHIGE, S. and SAUER, B. (1992). Genomic targetting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *PROC. NATL. ACAD. SCI. USA* **89**:7905-7909.
- FYNAN, E.F., WEBSTER, R.G., FULLER, D.H., HAYNES, J.R., et al. (1993). DNA vaccines: Protective immunizations by parental, mucosal, and gene-gun inoculations. *PROC. NATL. ACAD. SCI. USA* **90**:11478-11482.
- GALINSKI, M.S. (1991). Paramyxoviridae-transcription and replication. *ADVANCES IN VIRUS RESEARCH* **39**:129-162.
- GAUT, J.R. and HENDERSHOT, L.M. (1993). The modification and assembly of proteins in the endoplasmic reticulum. *CURR. OPIN. CELL BIOL.* **5**:589-595.

- GEISOW, M.J. (1992). *Glycoprotein glycans-roles and controls*. TIBTECH 10:,333-335.
- GELLISSEN, G. (1992). *Progress in developing methylotrophic yeasts as expression systems*. TIBTECH 10:,413-417.
- GERARD, R.D. and GLUZMAN, Y. (1985). *New host cell system for regulated simian virus 40 DNA replication*. MOL. AND CELL. BIOL. 5(11):,3231-3240.
- GETHING, M.J., MCCANNON, K. and SAMBROOK, J. (1986). *Expression of wild-type and mutant forms of influenza virus hemagglutinin: the role of folding intracellular transport*. CELL 46:,938-952.
- GETHING, M-J. (1991). *Molecular chaperones: individualists or groupies*. CURR. OPIN. CELL BIOL. 3:,610-614.
- GETHING, M-J. and SAMBROOK, J. (1992). *Protein folding in the cell*. NATURE 355:,33-44.
- GHOSH, G., VANDUYNE, G., GHOSH, S. and SIGLER, P.B. (1995). *Structure of NF- κ B p50 homodimer bound to a κ B site*. NATURE 373:,303-310.
- GIANNI, M., GALLINA, A., BARBANTI-BRODANO, G. and MILANESI, G. (1988). *Constitutive and inducible expression of cloned human interferon- β gene in hela cells through an episomal eukaryotic vector*. BBRC 157(3):,1175-1181.
- GILBOA, E. and SMITH, C. (1994). *Gene therapy for infectious diseases: the AIDS model*. TIG 10(4):,139-144.
- GILMORE, R. (1991). *The protein translocation apparatus of the rough endoplasmic reticulum, its associated proteins, and the mechanism of translocation*. CURR. OPIN. CELL BIOL. 3:,580-584.
- GLORIOSO, J.C., DELUCA, N.A. and FINK, D.J. (1995). *Development and application of herpes simplex virus vectors for human gene therapy*. ANNULA REVIEWS IN MICROBIOLOGY 49:,675-710.
- GLUZMAN, Y. (1981). *SV40-transformed simian cells support the replication of early SV40 mutants*. CELL 23:,175.
- GOEDDEL, D.V. (1990). *Gene expression technology*. METHODS IN ENZYMOLOGY 185:,1-471.
- GOEDDEL, D.V. (1990). *Systems for heterologous gene expression*. METHODS IN ENZYMOLOGY 185:,3-7.
- GOLDSTEIN, S., FORDIS, C.M. and HOWARD, B.H. (1989). *Enhanced transfection efficiency and improved cell-survival after electroporation of G2/M synchronised cells and treatment with sodium butyrate*. NAR 17(10):,3959-3971.
- GOODBURN, S. *Eukaryotic gene transcription*, Oxford:IRL PRESS, 1996.
- GOSSEN, M. and BUJARD, H. (1992). *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters*. PROC. NATL. ACAD. SCI. USA 89:,5547-5551.
- GOSSEN, M., BONIN, A.L., FREUNDLIEB, S. and BUJARD, H. (1994). *Inducible gene expression systems for higher eukaryotic cells*. CURR. OPIN. BIOTECH. 5:,516-520.
- GOSSEN, M., FREUNDLIEB, S., BENDER, G., MULLER, G., et al. (1995). *Transcriptional activation by tetracyclines in mammalian cells*. SCIENCE 268:,1766-1769.

GRAHAM, F.L. AND PREVEC, L. Adenovirus based expression vectors and recombinant vaccines. In: *Vaccines: New approaches to immunological problems*, edited by Ellis, R.W. Butterworth, 1992, p. 363-390.

GRAHAM, F.L., SMILEY, J., RUSSELL, W.C. and NAIRU, R. (1977). *Characteristics of a human cell-line transformed by DNA from human adenovirus type 5*. J. GEN. VIROL. 36:59-72.

GRUBER, C. and LEVINE, S. (1983). *Respiratory syncytial virus polypeptides.3. The envelope-associated proteins*. J. GEN. VIROL. 64:825-832.

GRUBER, C. and LEVINE, S. (1985). *Respiratory syncytial virus polypeptides.4. The oligosaccharides of the glycoproteins*. J. GEN. VIROL. 66:417-432.

GRUNERT, S. and JACKSON, R.J. (1994). *The immediate downstream codon strongly influences the efficiency of utilisation of eukaryotic translation initiation codons*. EMBO J 13(15):3618-3630.

GUPTA, R.K., RELYVELD, E.H., LINDBLAD, E.B., BIZZINI, B., et al. (1993). *Adjuvants-a balance between toxicity and adjuvanticity*. VACCINE 11:293-306.

HAASE, S.B., HEINZEL, S.S. and CALOS, M.P. (1994). *Transcription inhibits the replication of autonomously replicating plasmids in human cells*. MOL. AND CELL. BIOL. 14(4):2516-2524.

HAASE, S.B., HEINZEL, S.S., KRYSAN, P.J. and CALOS, M.P. (1989). *Improved EBV shuttle vectors*. MUTATION RESEARCH 220:125-132.

HAMAGUCHI, M., YOSHIDA, T., NISHIKAWA, K., NARUSE, H., et al. (1983). *Transcriptive complex of Newcastle disease virus I. Both L and P proteins are required to constitute an active complex*. VIROLOGY 128:105-117.

HAMMOND, C., BRAAKMAN, I. and HELENIUS, A. (1994). *Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality-control*. PROC. NATL. ACAD. SCI. USA 91(3):913-917.

HAMMOND, C. and HELENIUS, A. (1995). *Folding of VSV G-protein - sequential interaction with Bip and calnexin*. SCIENCE 266(5184):456-458.

HANKE, T. and RANDALL, R.E. (1994). *Processing of viral proteins for presentation by molecules of the major histocompatibility complex*. REV. MED. VIROL. 4:47-61.

HANKE, T., SZAWLOWSKI, P. and RANDALL, R.E. (1992). *Construction of solid matrix-antibody-antigen complexes containing simian immunodeficiency virus p27 using tag-specific monoclonal antibody and tag-linked antigen*. J. GEN. VIROL. 73:653-660.

HANKE, T., YOUNG, D.F., DOYLE, C., JONES, I., et al. (1995). *Attachment of an oligopeptide epitope to the C-terminus of recombinant SIV gp160 facilitates the construction of SMAA complexes while preserving CD4 binding*. J VIROL METHODS 53:149-156.

HARRISON, T.M., HUDSON, K., MUNSON, S.E., UFF, S., et al. (1995). *Derivation and partial analysis of two highly active myeloma cell transfectants*. BIOCHIMICA ET BIOPHYSICA ACTA 1260:147-156.

HASS, J., PARK, E.C. and SEED, B. (1996). *Codon usage limitation in the expression of HIV-1 envelope glycoprotein*. CURRENT BIOLOGY 6(3):315-324.

HEINZEL, S.S., KRYSAN, P.J., CALOS, M.P. and DUBRIDGE, R.B. (1988). *Use of simian virus 40 replication to amplify Epstein-Barr virus shuttle vectors in human cells*. J. VIROL. 62(10):3738-3746.

- HERRLER, G. and COMPANS, R.W. (1983). *Posttranslational modification and intracellular transport of mumps virus glycoproteins*. J. VIROL. 47:354-362.
- HIEBERT, S.W., PATERSON, R.G. and LAMB, R.A. (1985). *Haemagglutinin-neuraminidase protein of the paramyxovirus simian virus 5: nucleotide sequence of the mRNA predicts an N-terminal anchor*. J. VIROL. 54:1-6.
- HIEBERT, S.W., PATERSON, R.G. and LAMB, R.A. (1988). *Cell surface expression and orientation in membranes of the 44 amino acid SH protein of simian virus 5*. J. VIROL. 62:2347-2357.
- HIGH, S. (1995). *Protein translocation at the membrane of the endoplasmic reticulum*. PROG BIOPHYS MOLEC BIOL 63:233-250.
- HIRANO, A. and WONG, T.C. (1991). *Generation of mammalian-cells expressing stably measles-virus proteins via bicistronic RNA*. J VIROL METHODS 33:135-147.
- HIRT, B. (1967). *Selective extraction of polyoma DNA from infected mouse cell cultures*. J. MOL. BIOL. 26:365-369.
- HISHIYAMA, M., TANABAYASHI, K., TAKEUCHI, K. and YAMADA, A. (1996). *Establishment of cell-lines stably expressing mumps-virus glycoproteins*. JAPANESE JOURNAL OF MEDICAL SCIENCE & BIOLOGY 49(1):29-38.
- HITT, M., BETT, A.J., ADDISON, C., PREVEC, L., et al. (1995). *Techniques for human adenovirus vector construction and characterisation*. METHODS IN MOLECULAR GENETICS 7:13-30.
- HOHFELD, J. and HARTL, F.U. (1994). *Post-translational protein import and folding*. CURR. OPIN. IN BIOTECH. 6:499-509.
- HORIKAMI, S.M., CURRAN, J. and KOLAKOSKY, D. (1992). *Complexes of sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vivo*. J. VIROL. 66:4901-4908.
- HORVATH, C.M. and LAMB, R.A. (1992). *Studies on the fusion peptide of a paramyxovirus fusion glycoprotein: roles of conserved residues in cell fusion*. J. VIROL. 66(4):2443-2455.
- HU, A.H., CATTANEO, R. and SCHWARTZ, S. (1994). *Role of N-linked oligosaccharide chains in the processing and antigenicity of measles virus haemagglutinin protein*. J. GEN. VIROL. 75:1043-1052.
- HU, A.H. and NORRBY, E. (1994). *Role of individual cysteine residues in the processing and antigenicity of the measles virus haemagglutinin protein*. J. GEN. VIROL. 75:2173-2181.
- IMLER, J-L. (1996). *Adenovirus vectors as recombinant viral vaccines*. VACCINE 13(13):1143-1151.
- IWATA, S., SCHMIDT, A., TITANI, K., SUZUKI, M., et al. (1994). *Assignment of disulfide bridges in the fusion glycoprotein of sendai virus*. J. VIROL. 68:3200-3206.
- JALANKO, A., KALLIO, A., SALMINEN, M. and ULMANEN, I. (1989). *Efficient synthesis of influenza virus haemagglutinin in mammalian cells with an extrachromosomal Epstein-Barr virus vector*. GENE 78:287-296.
- JANKNECHT, R. and NORDHEIM, A. (1992). *Affinity purification of histidine-tagged proteins transiently produced in Hela cells*. GENE 121:321-324.

- JIN, H., SUBBARAO, K., BAGAI, S., LESER, G.P., et al. (1996). *Palmitylation of the influenza virus hemagglutinin (H3) is not essential for virus assembly or infectivity*. J. VIROL. **70**(3):1406-1414.
- JUNGnickel, B. and Rapoport, T.A. (1995). *A posttargeting signal sequence recognition event in the endoplasmic-reticulum membrane*. CELL **82**:261-270.
- KANE, S.E. and Gottesman, M.M. (1993). *Use of multidrug resistance gene in mammalian expression vectors*. METHODS IN ENZYMOLOGY **217**:34-47.
- KAUFMAN, R.J. and Murtha, P. (1987). *Translational control mediated by eukaryotic initiation factor-2 is restricted to specific mRNAs in transfected cells*. MOL. AND CELL. BIOL. **7**(4):1568-1571.
- KAWAHARA, N., YANG, X.Z., SAKAGUCHI, T., KIYOTANI, K., et al. (1992). *Distribution and substrate specificity of intracellular proteolytic processing enzyme(s) for paramyxovirus fusion glycoproteins*. J. GEN. VIROL. **73**:583-590.
- KAWAMOTO, T., MAKINO, K., NIW, H., SUGIYAMA, H., et al. (1988). *Identification of the human beta-actin enhancer and its binding-factor*. MOL. AND CELL. BIOL. **8**:267-272.
- KEMBALL-COOK, G., GARNER, I., IMANAKA, Y., NISHIMURA, T., et al. (1994). *High-level production of human blood coagulation factors VII and XI using a new mammalian expression vector*. GENE **139**:275-279.
- KOHARA, A. and ET AL., (1992). *Conformation and length of the signal sequence affect processing of secretory protein*. FEBS **311**(3):226-230.
- KOZAK, M. (1989). *The scanning model for translation-an update*. J. CELL. BIOL. **108**(2):229-241.
- KRIEGLER, M. (1990). *Assembly of enhancers, promoters and splice signals to control expression of transferred genes*. METHODS IN ENZYMOLOGY **185**:512-527.
- KRONMAN, C., VELAN, B., GOZES, Y., LEITNER, M., et al. (1992). *Production and secretion of high levels of recombinant human acetylcholinesterase in cultured cell lines: microheterogeneity of the catalytic subunit*. GENE **121**:295-304.
- KRYSAN, P.J. and CALOS, M.P. (1993). *Epstein-Barr virus-based vectors that replicate in rodent cells*. GENE **136**:137-143.
- KUCHLER, K. (1993). *Unusual routes of protein secretion: the easy way out*. TRENDS IN CELL BIOLOGY **3**:421-425.
- LAMB, R.A. and CHOPPIN, P.W. (1977). *The synthesis of Sendai virus polypeptides in infected cells. III. Phosphorylation of polypeptides*. VIROLOGY **81**:382-397.
- LAMB, R.A. and KOLAKOSKY, D. *Paramyxoviridae: The viruses and their replication*. In: *Fields Virology*, edited by Fields, B.N., Knipe, D.M. and Howley, P.M. Philadelphia: Lippincott-Raven, 1996, p. 1177-1204.
- LAMB, R.A., MAHY, B.W.J. and CHOPPIN, P.W. (1976). *The synthesis of Sendai virus polypeptides in infected cells*. VIROLOGY **69**:116-131.
- LAURING, B., KREIBICH, G. and WIEDMANN, M. (1995). *The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent polypeptide associated complex*. PROC. NATL. ACAD. SCI. USA **92**(21):9435-9439.

LAURING, B., SAKAI, H., KREIBICH, G. and WIEDMANN, M. (1995). *Nascent polypeptide-associated complex protein prevents mistargeting of nascent chains to the endoplasmic reticulum*. PROC. NATL. ACAD. SCI. USA **92**(12):5411-5415.

LEBKOWSKI, J.S., MCNALLY, M.M., OKARMA, T.B. and LERCH, L.B. (1988). *Adeno-associated virus: a vector system for efficient introduction and integration of DNA into a variety of mammalian cell types*. MOL. AND CELL. BIOL. **8**:3988-3996.

LECHNER, J. (1989). *Structure and biosynthesis of prokaryotic glycoproteins*. ANNU. REV. BIOCHEM. **58**:173-194.

LEHMAN, D.J., ROOF, L.L., BRIDEAU, R.J., AEED, P.A., et al. (1993). *Comparison of soluble and secreted forms of human parainfluenza virus type-3 glycoproteins expressed from mammalian and insect cells as subunit vaccines*. J. GEN. VIROL. **74**(3):459-469.

LEIB, D.A. and OLIVO, P.D. (1993). *Gene delivery to neurons: Is herpes simplex virus the right tool for the job*. BIOESSAYS **15**(8):547-554.

LESER, G.P., ECTOR, K.J. and LAMB, R.A. (1996). *The paramyxovirus simian virus 5 hemagglutinin-neuraminidase glycoprotein, but not the fusion glycoprotein, is internalized via coated pits and enters the endocytic pathway*. MOL. BIOL. CELL **7**:155-172.

LETOURNEUR, F., GAYNER, E.C., HENNECKE, S., DEMOLLIERE, C., et al. (1994). *Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum*. CELL **79**:1199-1207.

LEVINSON, A.D. (1990). *Expression of heterologous genes in mammalian cells*. METHODS IN ENZYMOLOGY **185**:485-496.

LEVITSKAYA, J., CORAM, M., LEVITSKY, V., IMREH, S., et al. (1995). *Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen -1*. NATURE **375**:685-688.

LEVY, J.P., MULDOON, R.R., ZOLOTUKHIN, S. and LINK, C.J. (1996). *Retroviral transfer and expression of a humanised, red-shifted green fluorescent protein gene into human tumor cells*. NATURE BIOTECHNOLOGY **14**(5):610-614.

LEWIS, M.J. and PELHAM, H.R.B. (1992). *Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum*. CELL **68**:353-364.

LIEBER, A., SANDIG, V., SOMMER, W., BAHRING, S., et al. (1993). *Stable high-level gene expression in mammalian cells by T7 phage RNA polymerase*. METHODS IN ENZYMOLOGY **217**:47-66.

LILJESTROM, P. (1994). *Alphavirus expression systems*. CURR. OPIN. BIOTECH. **5**:485-500.

LILJESTROM, P. and GAROFF, H. (1991). *A new generation of animal cell expression vectors based on the Semliki forest virus replicon*. BIOTECH **9**:1356-1361.

LINGAPPA, V.R., CHAIDEZ, J., YOST, C.S. and HEDGPETH, J. (1984). *Determinants for protein localisation-beta-lactamase signal sequence directs globin across microsomal membranes*. PROC. NATL. ACAD. SCI. USA **81**(12):456-460.

LUPTON, S. and LEVINE, A.J. (1985). *Mapping genetic elements of Epstein-Barr virus that facilitate extrachromosomal persistence of Epstein-barr-virus derived plasmids in human cells*. MOL. AND CELL. BIOL. **5**:2533-2542.

- LUZIO, J.P. and BANTING, G. (1993). *Eukaryotic membrane traffic: retrieval and retention mechanisms to achieve organelle residence*. TIBS **18**;;395-398.
- LYKO, F., MARTOGLIO, B., JUNGnickel, B., RAPOPORT, T.A., et al. (1995). *Signal sequence processing in rough microsomes*. J. BIOL. CHEM. **270**(34);19873-19878.
- MACHAMER, C.E. (1993). *Targeting and retention of Golgi membrane proteins*. CURR. OPIN. CELL BIOL. **5**;;606-612.
- MACKETT, M., SMITH, G.L. and MOSS, B. (1982). *Vaccinia virus: A selectable eukaryotic cloning and expression vector*. PROC. NATL. ACAD. SCI. USA **79**;;7415-7419.
- MACKETT, M., SMITH, G.L. and MOSS, B. (1984). *General method for the production and selection of infectious vaccinia virus recombinants expressing foreign genes*. J. VIROL. **49**;;857-864.
- MADISON, E.L. and BIRD, P. (1992). *A vector pSHT for the expression and secretion of protein domains in mammalian cells*. GENE **121**;;179-180.
- MALLABIABARRENA, A., JIMENEZ, M., RICO, M. and ALARCON, B. (1995). *A tyrosine-containing motif mediates ER retention of CD3-epsilon and adopts a helix-turn structure*. EMBO J **14**(10);2257-2268.
- MANCINI, M., DAVIS, H., TIOLLAIS, P. and MICHEL, M.L. (1996). *DNA-based immunization against the envelope proteins of the hepatitis-B virus*. JOURNAL OF BIOTECHNOLOGY **44**(1-3);47-57.
- MARGOLSKEE, R.F., KAVATHAS, P. and BERG, P. (1988). *Epstein-Barr virus shuttle vector for stable episomal replication of cDNA expression libraries in human cells*. MOL. AND CELL. BIOL. **8**(7);2837-2847.
- MARKWELL, M.A.K. and FOX, C.F. (1980). *Protein-protein interactions within Paramyxoviruses identified by native disulphide bonding or reversible chemical cross linking*. J. VIROL. **33**;;152-166.
- MATTHEWS, R.E.F. (1982). *Classification and nomenclature of viruses*. INTERVIROLOGY **17**;;1-18.
- MELLITS, K.H. and MATHEWS, M.B. (1988). *Effects of mutations in stem and loop regions on the structure and function of adenovirus VA RNA₁*. EMBO J **7**(9);2849-2859.
- MIGLIACCIO, G., NICCHITTA, C.V. and BLOBEL, G. (1992). *The signal sequence receptor, unlike the signal recognition particle receptor, is not essential for protein translocation*. J. CELL. BIOL. **117**;;15-25.
- MIGNATTI, P., MORIMOTO, T. and RIFKIN, D.B. (1992). *Basic fibroblast growth-factor, a protein devoid of secretory signal sequence, is released by cells via a pathway independent of the endoplasmic-reticulum Golgi-complex*. JOURNAL OF CELLULAR PHYSIOLOGY **151**(1);81-93.
- MILSTEIN, C., BROWNLEE, G., HARRISON, T. and MATHEWS, M.B. (1971). *A possible precursor of immunoglobulin light chains*. NATURE NEW BIOLOGY **239**;;117-120.
- MINER, J.N. and HRUBY, D.E. (1990). *Vaccinia virus: a versatile tool for molecular biologists*. TIBTECH **8**;;20-25.
- MIZUSHIMA, S. and NAGATA, S. (1990). *pEF-BOS, a powerful mammalian expression vector*. NAR **18**(17);5322.
- MORGAN, E.M., RE, G.G. and KINGSBURY, D.W. (1984). *Complete sequence of the Sendai virus NP gene from a cloned insert*. VIROLOGY **135**;;279-287.

MORRISON, T.G. (1988). *Structure, function and intracellular processing of paramyxovirus membrane-proteins*. VIRUS RESEARCH 10:113-135.

MORRISON, T.G. and PORTNER, A. Structure, function and intracellular processing of the glycoproteins of Paramyxoviridae. In: *The Paramyxoviruses*, edited by Kingsbury, .. New York: Plenum Publishing Corporation, 1991, p. 347-382.

MORRISON, T.G., MCQUAIN, C., OCONNELL, K.F. and MCGINNIS, L.W. (1990). *Mature cell associated HN protein of NDV exists in two forms differentiated by posttranslational modification*. VIRUS RESEARCH 5:113-134.

MORRISON, T.G. AND SIMPSON, D. Conformation and activity of the Newcastle disease virus HN protein in the absence of glycosylation. In: *The replication of negative strand viruses*, edited by Bishop, D.H.L. and Compans, R.H. Amsterdam: Elsevier/North Holland, 1981, p. 471-477.

MULLER, R. (1995). *Transcriptional regulation during the mammalian cell cycle*. TIG 11(5):173-178.

NAKAMURA, K., HOMMA, M. and COMPANS, R. (1982). *Effect of tunicamycin on the replication of Sendai virus*. VIROLOGY 119:447-487.

NANICHE, D., VARIOR-KRISHNAN, G., CERVONI, F., WILD, F., et al. (1993). *Human membrane cofactor protein (CD46) acts as a receptor for measles virus*. J. VIROL. 67:6025-6032.

NEUBERGER, M.S. and WILLIAMS, G.T. (1988). *The intron requirement for immunoglobulin gene expression is dependent upon the promoter*. NAR 16(14):6713-6724.

NG, D.T.W., HIEBERT, S.W. and LAMB, R.A. (1990). *Different roles of individual N-linked oligosaccharide chains in folding, assembly and transport of the simian virus 5 haemagglutinin-neuraminidase*. MOL. AND CELL. BIOL. 10(5):1989-2001.

NG, D.T.W., RANDALL, R.E. and LAMB, R.A. (1989). *Intracellular maturation and transport of the SV5 type II glycoprotein hemagglutinin-neuraminidase: specific and transient association with GRP78-BiP in the endoplasmic reticulum and extensive internalisation from the cell surface*. J. CELL. BIOL. 109(6 pt2):3273-3289.

NG, D.T.W. and WALTER, P. (1994). *Protein translocation across the endoplasmic reticulum*. CURR. OPIN. IN BIOTECH. 6:510-516.

NILSSON, T. and WARREN, G. (1994). *Retention and retrieval in the endoplasmic reticulum and Golgi apparatus*. CURR. OPIN. IN BIOTECH. 6:517-521.

NISHIO, M., TSURUDOME, M., KOMADA, H., KAWANO, M., et al. (1994). *Fusion properties of cells constitutively expressing human parainfluenza virus type 4A hemagglutinin-neuraminidase and fusion glycoproteins*. J. GEN. VIROL. 75(12):3517-3523.

NOVICK, S.L. and HOEKSTRA, D. (1988). *Membrane penetration of sendai virus glycoproteins during the early stage of fusion with liposomes as determined by hydrophobic affinity labelling*. PROC. NATL. ACAD. SCI. USA 85:7433-7437.

OGONAH, O.W., FREEDMAN, R.B., JENKINS, N., PATEL, K., et al. (1996). *Isolation and characterisation of an insect cell line able to perform complex N-linked glycosylation on recombinant proteins*. BIOTECHNOLOGY 14:197-202.

PANICALI, D. and PAOLETTI, E. (1981). Construction of poxviruses as cloning vectors: Insertion of the thymidine kinase gene from herpes simplex into the DNA of infectious vaccinia virus. *PROC. NATL. ACAD. SCI. USA* **79**:4927-4931.

PARKS, G. and LAMB, R.A. (1990). Defective assembly and transport of mutant paramyxovirus haemagglutinin-neuraminidase proteins containing altered cytoplasmic domains. *J. VIROL.* **64**(8):3605-3616.

PARKS, G. and LAMB, R.A. (1990). Folding and oligomerisation properties of a soluble and secreted form of the paramyxovirus haemagglutinin-neuraminidase glycoprotein. *VIROLOGY* **178**:498-508.

PARKS, G. and LAMB, R.A. (1991). Topology of eukaryotic type II membrane proteins: importance of N-terminal positively charged residues flanking the hydrophobic domain. *CELL* **64**:777-787.

PARKS, G.D. (1994). Mapping of a region of the paramyxovirus L protein required for the formation of a stable complex with the viral phosphoprotein P. *J. VIROL.* **68**:4862-4872.

PATERSON, R.G., HARRIS, T.J.R. and LAMB, R.A. (1984). Fusion protein of the paramyxovirus simian virus 5: nucleotide sequence predicts a highly hydrophobic glycoprotein. *PROC. NATL. ACAD. SCI. USA* **81**:6706-6710.

PATERSON, R.G., HIEBERT, S.W. and LAMB, R.A. (1985). Expression at the cell surface of biologically active fusion and Haemagglutinin-neuraminidase proteins of the paramyxovirus simian virus 5 from cloned cDNA. *PROC. NATL. ACAD. SCI. USA* **82**:7520-7524.

PATERSON, R.G. and LAMB, L.A. (1990). Conversion of a class II integral membrane protein into a soluble and efficiently secreted protein: multiple intracellular and extracellular oligomeric and conformational forms. *J. CELL. BIOL.* **110**:999-1011.

PATERSON, R.G. and LAMB, R.A. (1987). Ability of the hydrophobic fusion-related external domain of a paramyxovirus F protein to act as a membrane anchor. *CELL* **48**:441-452.

PATERSON, R.G., LAMB, R.A., MOSS, B. and MURPHY, B.R. (1987). Comparison of the relative roles of the F and HN surface glycoproteins of the paramyxovirus simian virus 5 in inducing protective immunity. *J. VIROL.* **61**(6):1972-1977.

PELHAM, H. (1992). The Florey lecture 1992, the secretion of proteins by cells. *PROC. R. SOC. LOND.* **250**:1-10.

PELHAM, H. (1994). About turn for COPS. *CELL* **79**:1125-1127.

PELHAM, H.R.B. (1990). The retention signal for soluble proteins of the endoplasmic reticulum. *TIBS* **15**:483-486.

PELHAM, H.R.B. (1991). Recycling of proteins between the endoplasmic reticulum and Golgi complex. *Curr. Opin. Cell Biol.* **3**:585-591.

PETITCLERC, D., ATTAL, J., THERON, M.C., BEARZOTTI, M., et al. (1995). The effect of various introns and transcription terminators on the efficiency of expression vectors in various cultured-cell lines and in the mammary-gland of transgenic mice. *JOURNAL OF BIOTECHNOLOGY* **40**(3):169-178.

POCH, O., BLUMBERG, B.M., BOUGUELERET, L. and TORDO, N. (1990). Sequence comparison of five polymerases (L proteins) of unsegmented negative strand RNA viruses: theoretical assignment of functional domains. *J. GEN. VIROL.* **71**:1153-1162.

PORTNER, A., SCROGGS, R.A. and METZGER, D.W. (1987). *Distinct functions of antigenic sites of the HN glycoprotein of sendai virus*. VIROLOGY 198:61-68.

PRECIOUS, B., YOUNG, D.F., BERMINGHAM, A., FEARN, R., et al. (1995). *Inducible expression of the P, V and NP genes of the paramyxovirus simian virus 5 in cell lines and an examination of NP-P and NP-V interactions*. J. VIROL. 69:8001-8010.

RANDALL, R.E., YOUNG, D.F., GOSWAMI, K.K.A. and RUSSELL, W.C. (1987). *Isolation and characterisation of monoclonal antibodies to simian virus 5 and their use in revealing antigenic differences between human, canine and simian isolates*. J. GEN. VIROL. 68:2769-2780.

RANDALL, R.E. and YOUNG, D.F. (1988). *Humoral and cytotoxic T cell immune responses to internal and external structural proteins of simian virus 5 induced by immunisation with solid matrix-antibody-antigen complexes*. J. GEN. VIROL. 69:2505-2516.

RANDALL, R.E., YOUNG, D.F. and SOUTHERN, J.A. (1988). *Immunisation with solid matrix-antibody-antigen complexes containing surface or internal virus structural proteins protects mice from infection with the paramyxovirus, simian virus 5*. J. GEN. VIROL. 69:2517-2526.

RANDALL, R.E. (1989). *Solid matrix-antibody-antigen (SMAA) complexes for constructing multivalent subunit vaccines*. IMMUNOLOGY TODAY 10:336-339.

RANDALL, R.E. and YOUNG, D.F. (1991). *Solid matrix-antibody-antigen complexes induce antigen specific CD8⁺ cells that clear a persistent paramyxovirus infection*. J. VIROL. 65(2):719-726.

RANDALL, R.E., HANKE, T., YOUNG, D.F. and SOUTHERN, J.A. (1993). *Two-tag purification of recombinant proteins for the construction of solid matrix-antibody-antigen (SMAA) complexes as vaccines*. VACCINE 11(12):1247-1252.

RANDALL, R.E. and SOUBERBIELE, B.E. *Presentation of virus antigens for the induction of protective immunity*. In: , 1994,

RANDALL, R.E., YOUNG, D.F., HANKE, T., SZAWLOWSKI, P., et al. (1994). *Purification of antibody-antigen complexes containing recombinant SIV proteins: comparison of antigen and antibody-antigen complexes for immune priming*. VACCINE 12(4):351-358.

RAPIEJKO, P.J. and GILMORE, R. (1994). *Signal sequence recognition and targeting of ribosomes to the endoplasmic-reticulum by the signal recognition particle do not require GTP*. MOL. BIOL. CELL 5(8):887-897.

REFF, M. (1993). *High-level production of recombinant immunoglobulins in mammalian cells*. CURR. OPIN. IN BIOTECH. 4:573-576.

REHEMTULLA, A. and KAUFMAN, R.J. (1992). *Protein processing within the secretory pathway*. CURR. OPIN. BIOTECH. 3:560-565.

RHODES, A.D., SPITALI, M., HUTCHINSON, G., RUD, E.W., et al. (1994). *Expression, characterisation and purification of simian immunodeficiency virus soluble, oligomerised gp 160 from mammalian cells*. J. GEN. VIROL. 75:207-213.

RICHARDSON, J.H. and MARASCO, W.A. (1995). *Intracellular antibodies: development and therapeutic potential*. TIBTECH 13:306-310.

ROBERTSON, J.S. (1994). *Safety considerations for nucleic acid vaccines*. VACCINE 12(16):1526-1528.

- ROSENBLUM, J.S., GILULA, N.B. and LERNER, R.A. (1996). On signal sequence polymorphisms and diseases of distribution. *PROC. NATL. ACAD. SCI. USA* **93**(9):4471-4473.
- ROTONDARO, L., MELE, A. and ROVERA, G. (1996). Efficiency of different viral promoters in directing gene expression in mammalian cells: effect of 3' - untranslated sequences. *GENE* **168**:195-198.
- ROUX, L., BEFFY, P. and PORTNER, A. (1985). Three variations in the cell surface expression of the haemagglutinin-neuraminidase glycoprotein of sendai virus. *J. GEN. VIROL.* **66**:987-1000.
- RUDERT, F., VISSER, E., GRADL, G., GRANDISON, P., et al. (1996). pLEF, a novel vector for expression of glutathione S-transferase fusion proteins in mammalian cells. *GENE* **169**:281-282.
- RUSSELL, R., PATERSON, R.G. and LAMB, R.A. (1994). Studies with cross-linking reagents on the oligomeric form of the paramyxovirus fusion protein. *VIROLOGY* **199**:160-168.
- RYAN, K.W. and PORTNER, A. (1990). Separate domains of Sendai virus P protein are required for binding to viral nucleocapsids. *VIROLOGY* **174**:515-521.
- SABBIONI, S., NEGRINI, M., RIMESSI, R., MANSERVIGI, R., et al. (1995). A BK virus episomal vector for constitutive high expression of exogenous cDNAs in human cells. *ARCH. VIROL.* **140**:335-339.
- SAMBROOK, J., FRITSCH, E.F. AND MANIATIS, T. Expression of cloned genes in cultures mammalian cells. In: *Molecular Cloning: A laboratory manual*, edited by Ford, N., Nolan, C. and Ferguson, M. New York: Cold Spring Harbor Laboratory Press, 1989, p. 16.3-16.29.
- SANDERSON, C.M., MCQUEEN, N.L. and NAYAK, D.P. (1993). Sendai virus assembly-m-protein binds to viral glycoproteins in transit through the secretory pathway. *J. VIROL.* **67**:651-663.
- SANDERSON, C.M., AVALOS, R., KUNDU, A. and NAYAK, D.P. (1995). Interaction of Sendai viral F, HN and M proteins with host cytoskeletal and lipid components in Sendai virus-infected cells. *VIROLOGY* **209**:701-707.
- SAUER, B. (1994). Site-specific recombination: developments and applications. *CURR. OPIN. BIOTECH.* **5**:521-527.
- SCHEID, A. and CHOPPIN, P.W. (1974). Identification of biological activities of paramyxovirus glycoproteins. Activation of cell fusion, hemolysis and infectivity by proteolytic cleavage of an inactive precursor protein of Sendai virus. *VIROLOGY* **57**:470-490.
- SERGEL, T. and MORRISON, T.G. (1995). Mutations in the cytoplasmic domain of the fusion glycoprotein of Newcastle disease virus depress syncytia formation. *VIROLOGY* **210**:264-272.
- SERGEL-GERMANO, T. (1994). Mutations in the fusion peptide and heptad repeat regions of the Newcastle disease virus fusion protein block fusion. *J. VIROL.* **68**(11):7654-7658.
- SHAW, A.S., ROTTIER, P.J.M. and ROSE, J.K. (1988). Evidence for the loop model of signal-sequence insertion into the endoplasmic-reticulum. *PROC. NATL. ACAD. SCI. USA* **85**(20):7592-7596.
- SHEN, E.S., COOKE, G.M. and HORLICK, R.A. (1995). Improved expression cloning using reporter genes and Epstein-Barr virus ori-containing vectors. *GENE* **156**:235-239.
- SHENK, T. Adenoviridae: The viruses and their replication. In: *Fields Virology*, edited by Fields, B.N., Knipe, D.M. and Howley, P.M. Philadelphia: Lippincott-Raven, 1996, p. 2111-2148.

SIMON, S. (1993). *Translocation of proteins across the endoplasmic reticulum*. CURR. OPIN. CELL BIOL. 5:581-588.

SIMPSON, D.A. and LAMB, R.A. (1992). *Alterations to influenza virus haemagglutinin cytoplasmic tail modulate virus infectivity*. J. VIROL. 66(2):790-803.

STARY, A. and SARASIN, A. (1992). *Simian virus 40 (SV40) large T antigen-dependent amplification of an Epstein-Barr virus-SV40 hybrid shuttle vector integrated into the human Hela cell genome*. J. GEN. VIROL. 73:1679-1685.

SUGDEN, B., MARSH, K. and YATES, J.L. (1985). *A vector that replicates as a plasmid and can be efficiently selected in B lymphocytes transformed by Epstein-Barr virus*. MOL. AND CELL. BIOL. 5:410-413.

TANIGAWA, T., HIKIDA, M., TAKAI, T. and OHMORI, H. (1993). *Enhancement of expression of introduced b-galactosidase gene by 2-aminopurine in mammalian cell lines*. JOURNAL OF FERMENTATION AND BIOENGINEERING 76(2):145-147.

TESHIGAWARA, T. and KATSURA, Y. (1992). *A simple and efficient mammalian gene expression system using an EBV-based vector transfected by electroporation in G2/M phase*. NAR 20(10):2607.

TIWARI, R.K. and ET AL, (1988). *GENE INDUCTION BY INTERFERONS AND DOUBLE-STRANDED RNA: SELECTIVE INHIBITION BY 2-AMINOPURINE*. MOL. AND CELL. BIOL. 8(10):4289-4294.

TRATSCHIN, J-D., MILLER, I.L., SMITH, M.G. and CARTER, B.J. (1985). *Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells*. MOL. AND CELL. BIOL. 5(11):3251-3260.

TRIMBLE, R.B. and VEROSTEK, M.F. (1995). *Glycoprotein oligosaccharide synthesis and processing in yeast*. TRENDS IN GLYCOSCIENCE AND GLYCOTECHNOLOGY 7(33):1-3.

ULMER, J.B., DECK, R.R., DEWITT, C.M., FRIEDMAN, A., et al. (1994). *Protective immunity by intramuscular injection of low doses of influenza virus DNA vaccines*. VACCINE 12(16):1541-1544.

VALENZUELA, P.A., MEDINA, A., RUTTER, W.J., AMMERER, G., et al. (1982). *Synthesis and assembly of hepatitis B virus surface antigen particles in yeast*. NATURE 298:347-350.

VARGHESE, J.N., WEBSTER, R.G., LAVER, W.G. and COLMAN, P.M. (1983). *Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution*. NATURE 303:35-40.

VIELICH, A. and ZIFF, E. (1985). *Adenovirus E1a proteins repress transcription from the SV40 early promoter*. CELL 40:705-716.

VEIT, M., SCHMIDT, M.F.G. and ROTT, R. (1989). *Different palmitoylation of paramyxovirus glycoproteins*. VIROLOGY 168:173-176.

WALTER, P. and BLOBEL, G. (1981). *Translocation of proteins across the endoplasmic reticulum. III. Signal recognition particle (SRP) causes signal sequence-dependent and site specific arrest of chain elongation that is released by microsomal membranes*. J. CELL. BIOL. 91:557-561.

WALTERS, M.C., FIERING, S., EIDEMILLER, J., MAGIS, W., et al. (1995). *Enhancers increase the probability but not the level of gene expression*. PROC. NATL. ACAD. SCI. USA 92:7125-7129.

WANG, L-F., YU, M., WHITE, J.R. and EATON, B.T. (1996). BTag: a novel six-residue epitope tag for surveillance and purification of recombinant proteins. *GENE* 169:53-58.

WARD, G.A., STOVER, C.K., MOSS, B. and FUERST, T.R. (1995). Stringent chemical and thermal regulation of recombinant gene expression by vaccinia virus vectors in mammalian cells. *PROC. NATL. ACAD. SCI. USA* 92:6773-6777.

WARREN, C.E. (1993). Glycosylation. *CURR. OPIN. IN BIOTECH.* 4:596-602.

WAXHAM, M.N. and WOLINSKY, J.S. (1986). A fusing mumps virus variant selected from a nonfusing parent with the neuraminidase inhibitor 2-deoxy-2,3-dihydro-N-acetylneuraminic acid. *VIROLOGY* 151:286-295.

WERTZ, G.W., COLLINS, P.L., YUNG, H.A., GRUBER, C., et al. (1985). Nucleotide-sequence of the G-protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane-protein. *PROC. NATL. ACAD. SCI. USA* 82:4075-4079.

WHITE, J.M., KIELIAN, M. and HELENIUS, A. (1983). Membrane fusion proteins of enveloped animal viruses. *QUARTERLY REVIEW OF BIOPHYSICS* 16:151-195.

WIEDMANN, B., SAKAI, H., DAVIS, T.A. and WIEDMANN, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. *NATURE* 370:434-440.

WILD, T.F., FAYOLLE, J., BEAUVERGER, P. and BUCKLAND, R. (1994). Measles virus fusion: role of the cysteine-rich region of the fusion glycoprotein. *J. VIROL.* 68(11):7546-7548.

WILKINSON, G.W.G. and AKRIGG, A. (1991). The cytomegalovirus major immediate early promoter and its use in eukaryotic expression systems. *ADV. GENE TECH.* 2:287-310.

WILSON, C., GILMORE, R. and MORRISON, T.G. (1987). Translation and membrane insertion of the haemagglutinin-neuraminidase glycoprotein of Newcastle disease virus. *MOL. AND CELL. BIOL.* 7(4):1386-1392.

WILSON, I.A., SKEHEL, J.J. and WILEY, D.C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *NATURE* 289:366-373.

WOLFE, S.L. *Molecular and cellular biology*, Belmont:Wadsworth Publishing Company, 1993.

WYATT, L.S., MOSS, B. and ROZENBLATT, S. (1995). Replication-deficient vaccinia virus encoding bacteriophage T7 RNA polymerase for transient gene expression in mammalian cells. *VIROLOGY* 210:202-205.

XIANG, Z.Q., SPITALNIK, S.L., CHENG, J., ERIKSON, J., et al. (1995). Immune responses to nucleic acid vaccines to rabies virus. *VIROLOGY* 209:569-579.

XU, Z.Z., KROUGLIAK, V., PREVEC, L., GRAHAM, F.L., et al. (1995). Investigation of promoter function in human and animal cells infected with human recombinant adenoviruses expressing rotavirus antigen VP7sc. *J. GEN. VIROL.* 76:1971-1980.

YOUNG, D.F. M.Sc. thesis, University of St. Andrews, 1991.

YOUNG, J.M., CHEADLE, C., FOULKE, J.S., DROHAN, W.N., et al. (1988). Utilisation of an Epstein-Barr virus replicon as a eukaryotic vector. *GENE* 62:171-185.

YATES, J.L., WARREN, N., REISMAN, D. and SUGDEN, B. (1984). A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. *PROC. NATL. ACAD. SCI. USA* 81:3806-3810.

YATES, J.L., WARREN, N. and SUGDEN, B. (1985). *Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells.* NATURE 313:812-815.

ZENTGRAF, H., FREY, M., SCHWINN, S., TESSMER, C., et al. (1995). *Detection of histidine-tagged fusion proteins by using a high-specific mouse monoclonal anti-histidine tag antibody.* NAR 23(16):3347-3348.

ZHANG, W.W., CHAREST, H. and MATLASHEWSKI, G. (1995). *The expression of biologically active human p53 in Leishmania cells: a novel eukaryotic system to produce recombinant proteins.* NAR 23(20):4073-4080.

ZOLOTUKHIN, S., POTTER, M., HAUSWIRTH, W.W., GUY, J., et al. (1996). *A humanised green fluorescent protein cDNA adapted for high-level expression in mammalian cells.* J. VIROL. 70(7):4646-4654.

ZWART, D.E., BREWER, C.B., LAZAROVITS, J., HENIS, Y.I., et al. (1996). *Degradation of mutant influenza-virus hemagglutinins is influenced by cytoplasmic sequences independent of internalization signals.* J. BIOL. CHEM. 271(2):907-917.